

VNIVERSITAT DE VALÈNCIA

Programa de Doctorado en Física



Tesis Doctoral

Implementación de un display 3D basado en imagen integral

Adrián Dorado Vide

Directores: Dr. Manuel Martínez Corral
Dr. Genaro Saavedra Tortosa

Septiembre 2018

D. Genaro SAAVEDRA TORTOSA y D. Manuel MARTÍNEZ CORRAL,
Catedráticos de Universidad del Departamento de Óptica y Optometría y
Ciencias de la Visión de la Universitat de València,

CERTIFICAN que la presente memoria, *Implementación de un display 3D
basado en imagen integral*, resume el trabajo de investigación realizado,
bajo su dirección, por D. Adrián DORADO VIDE y constituye su Tesis para
optar al Grado de Doctor en Física.

Y para que conste, en cumplimiento de la legislación vigente, firman
el presente certificado en Valencia, a 3 de septiembre de dos mil dieciocho.

Dr. Genaro Saavedra Tortosa

Dr. Manuel Martínez Corral

A mis padres

*Ni el más sabio conoce el fin de
todos los caminos.*

J. R. R. Tolkien, *El Señor de los
Anillos*

*Toda verdad es fácil de comprender
una vez que ha sido descubierta, el
problema es descubrirla.*

Galileo Galilei

Agradecimientos

A mis directores, Manuel Martínez Corral y Genaro Saavedra Tortosa. Tanto por su guía y consejos durante el desarrollo de esta tesis doctoral, como por su paciencia y comprensión ante mis inquietudes y preguntas.

Mi más sincero agradecimiento a los compañeros de fatigas del grupo *3D Imaging & Display Laboratory*, os echaré de menos. Con especial mención a Anabel Llavador, por su inestimable ayuda en la elaboración de esta Tesis.

Gracias a viejos amigos, como Héctor Vives, Sergio Gimeno y Javier Barrios por su apoyo y ayuda en estos años. Lo siento, no se me ocurre un lugar para meter la palabra Óptica.

Agradecer también a mi familia todo el apoyo que me habéis proporcionado.

Y, por último, gracias de todo corazón a ti, Isa, y no sólo por tu impagable ayuda en la escritura de esta Tesis. Sin ti, no sería la persona que soy ahora.

Resumen

Desde hace varios años, el desarrollo y la evolución de tecnologías capaces de registrar y/o reconstruir información 3D está en auge, principalmente debido al gran interés que despierta este campo de estudio. Entre estas tecnologías, destaca la técnica de *Imagen Integral*, debido principalmente a su capacidad de generar una imagen 3D que proporciona al sistema visual una respuesta similar a la que obtendría percibiendo un objeto 3D real. Aunque esta técnica no está exenta de limitaciones, se están realizando prometedores avances para resolver sus principales inconvenientes.

En esta Tesis se estudia la técnica de imagen integral con el objetivo de implementar un monitor 3D que sirva como prototipo para un sistema de captura y display de imágenes 3D en tiempo real. Para lograrlo, se ha diseñado un nuevo método que permite adquirir y adaptar la información 3D necesaria para producir una imagen 3D de calidad. Por otra parte, se han planteado dos aplicaciones para dicho método que lo hacen más versátil. En la primera, se aumenta el ángulo de visualización del monitor 3D, mientras que la segunda aplicación permite producir contenido 3D adicional a partir de la misma información. Para implementar el prototipo de monitor 3D, se ha utilizado el método propuesto junto con tecnología *eye-tracking*, lo que permite adaptar la información mostrada en pantalla con la posición del observador, mejorando así la experiencia de visualización.

Índice

Agradecimientos	IX
Resumen	XI
Listado de abreviaturas	XVII
1 Introducción	1
1.1 Objetivos	3
2 Imágenes 3D y la función plenóptica	5
2.1 La idea tras el monitor de InI	6
2.2 La función plenóptica	6
2.2.1 Función plenóptica de una escena 3D	9
2.2.2 Transformación de la función plenóptica	11
3 Captura de la función plenóptica	15
3.1 Cámara convencional	15
3.2 InI	18
3.2.1 Muestreo de la función plenóptica en un sistema de InI	23
3.3 Fotografía plenóptica	24
3.3.1 Muestreo de la función plenóptica en un sistema de fotografía plenóptica	32
3.4 Equivalencia entre InI y fotografía plenóptica	34
3.5 Casos reales	37
4 Reconstrucción de la función plenóptica	39

4.1	El monitor de InI	39
4.1.1	Configuración simétrica y asimétrica de un monitor de InI	41
4.1.2	Características de la imagen 3D generada	50
4.2	Visualización de la imagen 3D	51
4.2.1	Área de visualización de un monitor de InI	56
4.2.2	Resolución de la imagen percibida	60
4.2.3	Propuesta de monitor de InI	62
5	Método para la generación de microimágenes adaptadas a un monitor de InI	65
5.1	Adaptación de las microimágenes a un monitor de InI . . .	66
5.2	Sistemas de captura convencionales	73
5.3	Imágenes pseudoscópicas	74
5.4	SPOC 2.0	76
5.4.1	Regreso a la equivalencia entre InI y fotografía plenóptica	76
5.5	Resultados experimentales	81
6	Actualización del método SPOC 2.0	89
6.1	Aumento del ángulo de visualización de un monitor de InI	89
6.1.1	Ángulo de visualización y rotación de la MR	90
6.1.2	Combinación con SPOC 2.0	93
6.1.3	Resultados experimentales	96
6.2	Película 3D a partir de una única imagen integral	99
6.2.1	Generar secuencias de imágenes con SPOC 2.0 . .	100
6.2.2	Mapa de disparidad	101
6.2.3	Resultados experimentales.	102
7	Mejora de la experiencia de visualización de un monitor de InI mediante un dispositivo eye-tracking	109
7.1	Aumento del área de visualización y paralaje mediante un dispositivo eye-tracking y el método SPOC 2.0.	110
7.1.1	Método	110

7.2 Resultados experimentales	114
8 Conclusiones	123
A Anexo	125
B Listado de vídeos	129
Bibliografía	131

Listado de abreviaturas

AS	Área de seguimiento
AV	Área de visualización
EI	Imagen elemental (<i>Elemental image</i>)
InI	Imagen integral (<i>Integral imaging</i>)
LC	Lente de la cámara
ML	Matriz de lentes
MR	Matriz de reconstrucción
PR	Plano de referencia
PRI	Plano de referencia imagen
PV	Plano de visualización
SAD	<i>Sum of Absolute Differences</i>

Introducción

En el año 1838, C. Wheatstone construyó el primer estereoscopio [1], un aparato capaz de producir imágenes con sensación de profundidad. Desde entonces, el interés en la captura y proyección de información 3D no ha hecho más que crecer. Un claro ejemplo son las películas 3D que podemos disfrutar en la actualidad, cuyo origen se remonta a los anaglifos propuestos por W. Rollmann en 1853 [2]. Aunque ahora se utilizan técnicas más modernas (en lugar de usar gafas con filtros de colores se utilizan gafas con cristales polarizados [3] u obturadores electrónicos [4]) las técnicas que se emplean en el cine o televisión 3D, denominadas estereoscópicas, se basan en el mismo principio: inducir disparidad binocular al proporcionar una perspectiva diferente a cada ojo del observador, lo que da lugar a una sensación de profundidad. Hay que destacar que también existen sistemas estereoscópicos que no necesitan de la ayuda de ningún elemento externo, como gafas especiales, y que en su lugar utilizan barreras de paralaje u hojas lenticulares [5], [6]. En este caso, el sistema se conoce como autoestereoscópico, y suelen ser más cómodos para el observador. No obstante, todos los sistemas estereoscópicos presentan dos inconvenientes importantes: sólo proporcionan una perspectiva de la escena y su utilización causa fatiga visual debido al conflicto convergencia-acomodación [7]-[10].

Otro tipo de sistemas de imagen 3D más complejos que los estereoscópicos, son aquellos que proporcionan al observador múltiples perspectivas de la escena como, por ejemplo, la imagen integral [11], [12] (InI, del inglés *integral imaging*), la holografía [13]-[15] o los sistemas lenticulares [13], [16]-[18]. Lejos de comparar o analizar las distintas técnicas de ima-

gen 3D existentes, en esta Tesis estamos interesados en la que, en nuestra opinión, es la más prometedora para la creación de contenido 3D, la InI. Esta técnica fue propuesta por Lippmann en 1908 [11], bajo el nombre de *Fotografía Integral*. En el método original propuesto por Lippmann, una matriz de estenopes o lentes forma un conjunto de perspectivas diferentes de un objeto sobre una emulsión fotográfica. Tras revelar la película fotográfica, ésta se sitúa tras la misma matriz de estenopes o lentes que sirvió para grabarla. Al iluminar el sistema con luz difusa, por el principio de reversibilidad de los rayos, se reconstruye una imagen 3D (mediante el cruce de pinceles de luz) que puede ser visualizada por un observador situado delante del sistema de fotografía integral. Debido a la mala calidad de las lentes de la época en la que apareció la fotografía integral, la técnica no tuvo mucha aceptación y no fue hasta que se consiguió mejorar el proceso de fabricación de las lentes cuando creció el interés por ella. Entre los avances realizados en la técnica de InI, destacan las propuestas de Ives y Coffey [19], [20], que introdujeron una lente de campo durante la captura para formar la imagen sobre la matriz de lentes, y la posterior implementación de dicho sistema por parte de Adelson y Wang [21]. Asimismo, en 1997, Okano [22] actualizó la técnica de fotografía integral al emplear por primera vez dispositivos digitales y conseguir realizar una captura en tiempo real.

La InI presenta una serie de ventajas que han hecho que, en los últimos años, haya experimentado un rápido desarrollo. Entre estas ventajas destacan que, para su funcionamiento, no se requiere del uso de ningún elemento externo (como gafas especiales), puede utilizar iluminación incoherente (a diferencia de la holografía) o que proporciona paralaje horizontal y vertical de forma continua. Asimismo, como esta técnica produce una imagen 3D capaz de proporcionar al sistema visual la sensación de estar percibiendo un objeto tridimensional real, no produce el conflicto convergencia-acomodación. Debido a ello, se han producido múltiples estudios centrados tanto en conseguir solventar sus principales inconvenientes (como su limitado ángulo de visualización, baja resolu-

ción y la dificultad de implementar la técnica en tiempo real) como en encontrar nuevas aplicaciones. Entre estas últimas, destacan el reenfoque digital, eliminar oclusiones entre objetos, reconocimiento de objetos, la microscopia de InI, la obtención de mapas de profundidad o el análisis de tensiones en láminas con oclusiones, entre otras [12], [23]-[35]. No obstante, la generación de imágenes 3D mediante un monitor de InI sigue siendo una de sus aplicaciones más interesantes [36]-[39].

1.1 Objetivos

El objetivo principal de esta Tesis es, como indica el título de la misma, la implementación de un display 3D basado en InI. En concreto, diseñaremos un monitor de InI capaz de producir imágenes 3D a partir de dispositivos comerciales. No obstante, tanto en el proceso de captura de la información 3D, como en el proceso de reconstrucción, utilizaremos componentes reales. Es decir, la imagen 3D se formará a partir de la información capturada de una escena 3D real, obtenida mediante el uso de una cámara digital, en lugar de realizar una captura sintética por ordenador de una escena virtual. Asimismo, nuestra intención es que este dispositivo sirva como prototipo para un sistema de captura y display de imágenes 3D en tiempo real que sea capaz de proporcionar una buena experiencia de visualización al observador. Para ello, analizaremos las características de los sistemas de InI, lo que nos permitirá diseñar un dispositivo que genere imágenes 3D de calidad. Por otro lado, propondremos y mejoraremos un método que permita obtener la información 3D que se mostrará en el monitor de InI y que, también, permita mejorar la experiencia de visualización de la imagen 3D.

La estructura de esta Tesis se organiza de la siguiente manera. En el capítulo 2 se explicará la idea fundamental tras un monitor de InI y se analizará el concepto de la función plenóptica, lo que asentará las bases que nos permitirán describir las técnicas de captura de InI y fotografía

plenóptica (capaces de obtener información de la función plenóptica de una escena 3D) en el capítulo 3. En el capítulo 4 analizaremos en detalle las características y el funcionamiento de un monitor de InI para, al final, plantear las características del monitor de InI que queremos implementar. En el capítulo 5 propondremos un método para obtener la información 3D que será mostrada en el monitor de InI y que, además, permite solucionar algunos de los problemas que presentan estos sistemas. Lejos de abandonar este método, en el capítulo 6 presentaremos dos actualizaciones del mismo para, después, plantear un método que mejora la experiencia de visualización y el ángulo de visualización de un monitor de InI en el capítulo 7. Finalmente, en el capítulo 8 comentaremos las conclusiones de este trabajo.

Imágenes 3D y la función plenóptica

Como se ha comentado anteriormente, el objetivo principal de este trabajo es implementar un monitor de InI capaz de reproducir imágenes 3D. No obstante, antes de explicar las características y el funcionamiento de un monitor de InI, es conveniente comentar brevemente la idea que subyace detrás de este tipo de dispositivos y explicar su fundamento: la función plenóptica.

Es importante destacar que este trabajo se ha desarrollado dentro del marco de la aproximación paraxial de la óptica geométrica. En consecuencia, de aquí en adelante se considerará que la luz se propaga en forma de rayos, representados como líneas dibujadas en el espacio que corresponden a la dirección del flujo de energía radiante, descartando el comportamiento ondulatorio de la luz. Asimismo, en esta Tesis se considerará que los medios en los que se propagan los rayos son homogéneos y que se comportan del mismo modo en toda dirección (medios isótropos). Por tanto, los rayos son rectos y perpendiculares a los frentes de onda [40]. Finalmente, el criterio de signos que se seguirá es el siguiente: las distancias axiales se consideran positivas en el sentido de propagación de la luz y un ángulo será positivo si para llevar la recta que lo define al eje óptico ¹ del sistema, el recorrido más corto se realiza en sentido contrario a las agujas del reloj.

¹Línea recta que atraviesa el centro de simetría de los elementos del sistema óptico.

2.1 La idea tras el monitor de InI

Supongamos una serie de objetos 3D en un determinado volumen del espacio. Este conjunto de objetos, en esta Tesis denominado escena 3D o escena, o bien emiten luz o bien difunden la luz que reciben de una fuente externa. Por tanto, en los alrededores de dicha escena existe un flujo de energía luminosa, o flujo de rayos de luz, que es característico de cada escena 3D, ya que depende de los objetos que la componen. Por este motivo, el sistema visual humano reconocerá el flujo de luz procedente de una escena 3D particular como perteneciente a ésta de manera inequívoca.

Teniendo este último aspecto en cuenta, si se replicara el flujo luminoso proveniente de una escena 3D concreta y se enviara al sistema visual humano, éste interpretará el flujo de energía radiante como proveniente de dicha escena 3D. Así, el sistema visual percibirá una imagen 3D de la escena original sin estar necesariamente ante dicha escena. Justamente, éste es el fundamento de las técnicas de generación de imágenes 3D de InI basadas en las ideas propuestas por Lippmann [11]. Al reconstruir mediante un monitor de InI el flujo de luz procedente de una escena, se genera una imagen 3D que replica la experiencia que el sistema visual humano tendría al percibir la escena original.

A continuación, procederemos a modelizar el flujo de luz que proviene de una escena 3D. Con ello se creará la estructura necesaria que nos permitirá analizar el proceso mediante el cual el flujo de luz es capturado y reconstruido.

2.2 La función plenóptica

El flujo de rayos de luz originado por una escena 3D depende de los objetos que la componen. En concreto, depende de la radiancia, r ,

en cada punto de la superficie de los mismos. La radiancia tiene como unidad de medida $W\ sr^{-1}\ m^{-2}$ y se define como la cantidad de energía electromagnética que incide sobre, atraviesa o emerge de un elemento de superficie, dS , en un ángulo sólido de anchura $d\omega$ por unidad de tiempo [41]. Esta magnitud se puede expresar como:

$$r = \frac{d^2\Phi}{d\omega ds \cos \theta}, \quad (2.1)$$

donde θ representa el ángulo formado entre la normal a la superficie y la dirección estudiadas, y Φ representa el flujo radiante².

La función que describe el flujo de luz en términos de la radiancia se denomina función plenóptica [21], [42] o **light field** [43], [44]. La función plenóptica es una función 5D, $L(x, y, z, \theta, \varphi)$, que indica la radiancia de los rayos y su dirección (θ, φ) en cada punto del espacio 3D (x, y, z) , ver Figura 2.1. Por conveniencia, en lugar de considerar todos los rayos procedentes de la escena 3D, en esta Tesis sólo tendremos en cuenta los rayos que se propagan en la dirección de nuestro interés: la que coincidirá básicamente con el eje óptico de los sistemas ópticos empleados en la captura y reconstrucción de la función plenóptica, z en nuestro caso.

Asimismo, a lo largo de esta Tesis se asumirá que los medios que atraviesan los rayos de luz son no dispersivos y libres de elementos oclusores. En consecuencia, la radiancia de los rayos permanece inalterada durante su propagación, convirtiendo la coordenada z en redundante. Bajo estas condiciones la función plenóptica se reduce a 4D $L(x, y, \theta, \varphi)$ [44], [45]. Esta simplificación no solo facilita trabajar con la función plenóptica sino que, además, tiene una interesante consecuencia: una vez conocida la función plenóptica en un plano z_0 perpendicular al eje óptico, ésta queda completamente definida en el resto de planos paralelos

²Tasa de energía radiante que incide sobre, atraviesa o emerge de una superficie dada por unidad de tiempo.

al mismo. Esta característica es vital para los monitores de InI, ya que implica que es posible recuperar la función plenóptica original, emitida por una escena 3D, capturándola en cualquier plano perpendicular a su dirección principal de propagación. Además, permite realizar algunas de las aplicaciones mencionadas en la introducción, como el reenfoque digital [23], [24], [26].

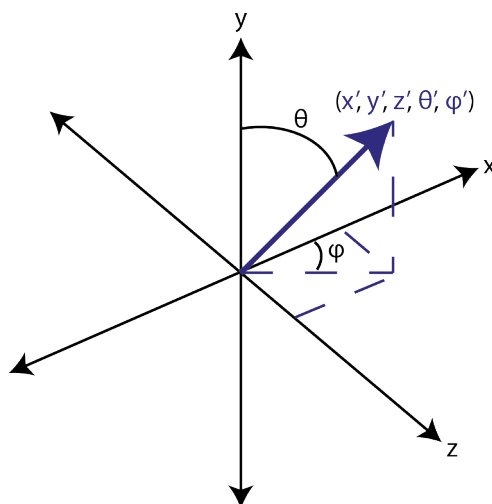


Figura 2.1.: Parametrización de un rayo de luz en la función plenóptica

Visualizar una función de 4D es complicado. Afortunadamente, es posible representar un corte 2D de la función plenóptica sin que esto suponga una pérdida de generalidad ni oculte los conceptos y propiedades más importantes. Para ello, basta con caracterizar los rayos de luz en función de una coordenada espacial, x , y una angular, θ , que indiquen su posición e inclinación cuando intersecten con un plano z perpendicular a la dirección de propagación, ver Figura 2.2 (a). Bajo esta parametrización, se puede representar cada rayo de la función plenóptica como un punto en un diagrama espacio-angular [46], [47], ver Figura 2.2 (b). Es conveniente mencionar que la representación en diagramas espacio-angulares es conocida también con el nombre de imágenes epipolares o EPI (del inglés *Epipolar images*) [44], [47]. Nótese que un diagrama espacio-angular de la función plenóptica representa un corte 2D de la misma, uno de los muchos

posibles de la función plenóptica $4D^3$. A pesar de que esta representación 2D se realiza por simplicidad, la extensión a 4D es inmediata.

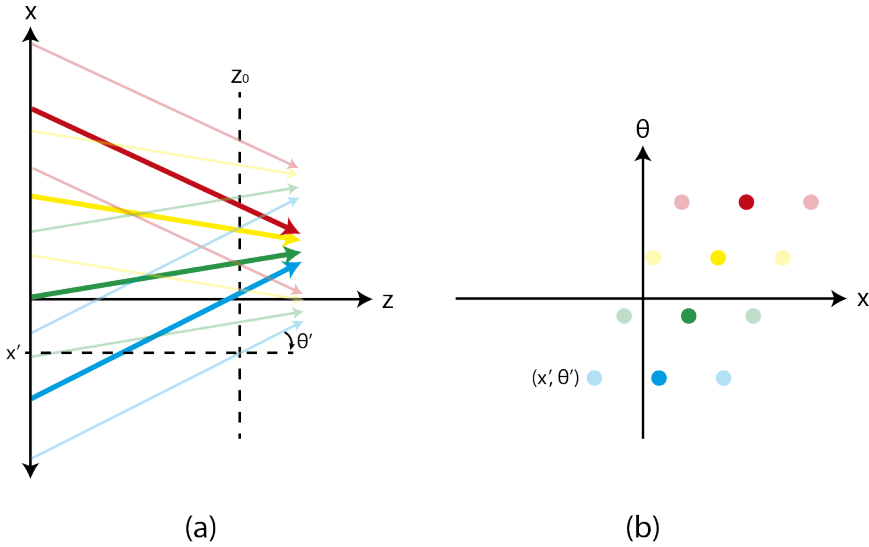


Figura 2.2.: (a) Parametrización de un rayo de la función plenóptica en su intersección con un plano z_0 perpendicular a la dirección de propagación. (b) Diagrama espacio-angular de la función plenóptica representada en (a) en el plano z_0 . Cada rayo aparece como un punto en el diagrama espacio-angular.

2.2.1 Función plenóptica de una escena 3D

Para complementar esta descripción de la función plenóptica, a continuación, se analiza la representación espacio-angular en dos casos específicos.

Comenzamos primero analizando una versión simplificada de una escena 3D, en la que sólo consideramos los rayos procedentes de dos puntos separados axial y lateralmente, ver Figura 2.3 (a). La representación espacio-angular de la función plenóptica, originada por los dos puntos

³El número de diagramas espacio-angulares (cortes 2D) que se pueden generar de la función plenóptica es considerablemente grande. En concreto, es igual al producto de las dimensiones de las cuatro variables de la función, (x, y, θ, φ) .

de la escena, en un plano perpendicular a la dirección de propagación a una distancia z_0 del objeto más próximo, se muestra en la Figura 2.3 (b). En dicha Figura se aprecia que los rayos procedentes de un mismo punto de la escena 3D se representan como una recta en el diagrama espacio-angular. Asimismo, la pendiente de dicha recta depende de la distancia del punto de origen de los rayos al plano en el que se analiza la función plenóptica.

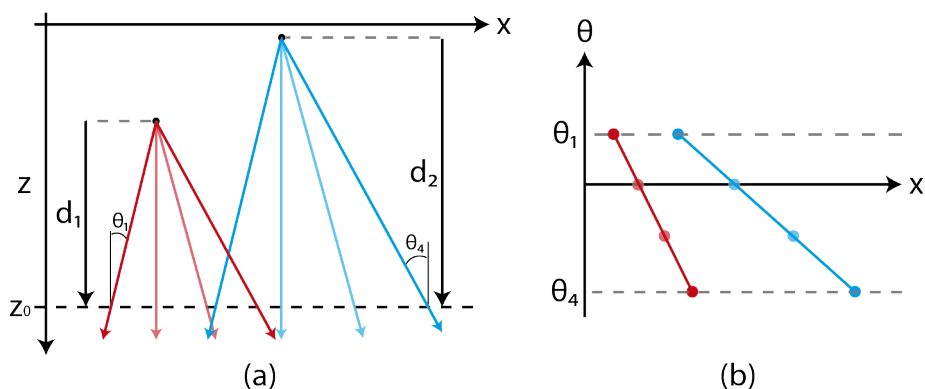


Figura 2.3.: (a) Flujo de rayos emitido por dos puntos separados axial y lateralmente. (b) Diagrama espacio-angular de la función plenóptica de la escena representada en (a) en un plano z_0 perpendicular a la dirección de propagación.

Ahora, repetiremos el análisis con una escena 3D completa, compuesta por dos objetos separados axial y lateralmente. En la Figura 2.4 se muestra la fotografía de la escena 3D y la representación espacio-angular (o EPI) de su función plenóptica en el plano del objetivo de la cámara que tomó la fotografía. Como en la versión simplificada, los rayos procedentes de un mismo punto aparecen como una línea en el diagrama espacio-angular. Por tanto, cada objeto de la escena 3D se muestra como un conjunto de líneas cuya inclinación depende de la distancia entre el objeto y el plano en el que se analiza la función plenóptica. En el ejemplo de la Figura 2.4, el optometrista se encuentra más cerca de la cámara que la figura de Einstein y, por tanto, la pendiente es mayor.

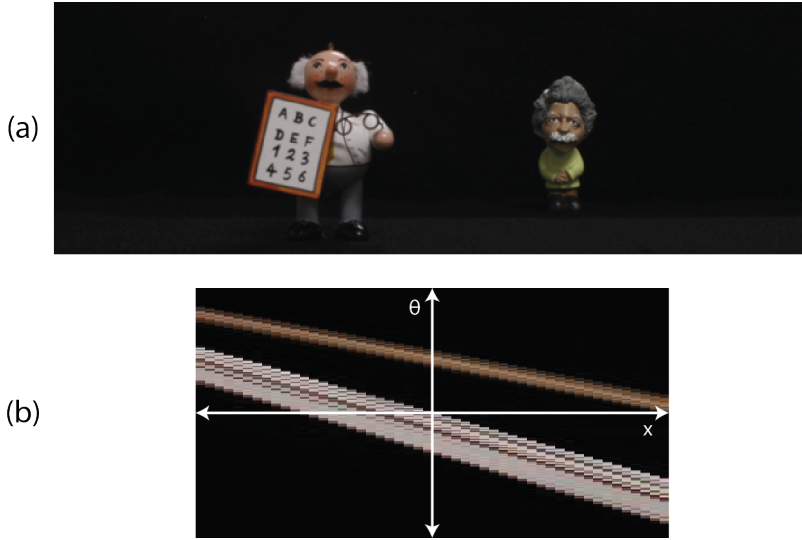


Figura 2.4.: (a) Fotografía de una escena 3D compuesta por dos objetos, situados a distinta distancia de la cámara. (b) Representación espacio-angular (o EPI) de la función plenóptica en el plano del objetivo de la cámara.

2.2.2 Transformación de la función plenóptica

Como hemos comentado anteriormente y en consonancia con el principio de conservación de la energía, la función plenóptica se conserva en su propagación. Es decir, la radiancia de cada uno de los rayos en un plano z_0 , perpendicular a la dirección de propagación, es la misma que en un plano z_1 posterior. Esto puede expresarse como:

$$L_0(x, \theta) = L_1(x', \theta'), \quad (2.2)$$

donde (x, θ) y (x', θ') representan las coordenadas espacio-angulares antes y después de la transformación, respectivamente. Por tanto, conocida la función plenóptica en un plano, es posible obtenerla en cualquier otro plano. Para ello, sólo es necesario saber qué transformaciones sufren los rayos de luz durante su propagación, ya sea cuando viajan libremente

o cuando se encuentran con elementos ópticos que no disipen energía. Para esta Tesis, sólo se considerarán las dos transformaciones que son necesarias para nuestro estudio: la propagación libre y el efecto de las lentes delgadas.

Bajo el marco de la óptica geométrica, en su aproximación paraxial, estas dos transformaciones son lineales y, por tanto, se pueden describir mediante matrices [48], [49], lo que facilita su análisis matemático, como veremos a continuación.

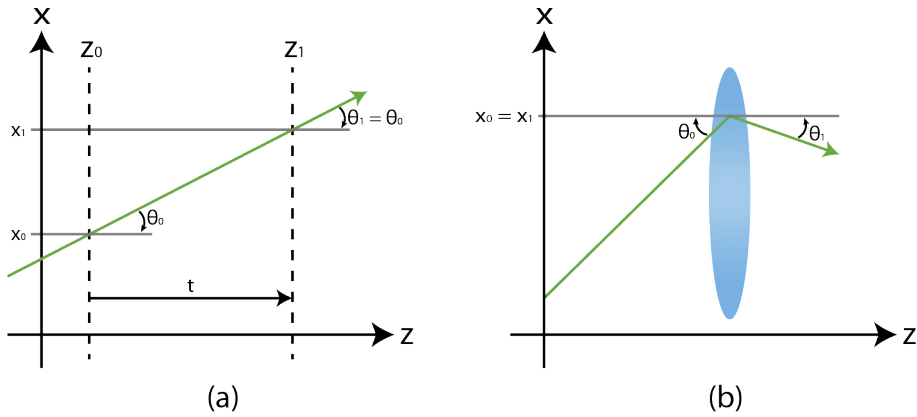


Figura 2.5.: (a) Alteración que sufre un rayo que se propaga libremente de un plano z_0 a un plano z_1 . (b) Alteración que sufre un rayo al atravesar una lente.

La matriz que describe la alteración que sufre un rayo de luz al propagarse libremente de un plano z_0 a un plano z_1 se expresa como:

$$T = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}, \quad (2.3)$$

donde t es la distancia que separa dichos planos. Por tanto, la transformación sufrida por los rayos durante la propagación libre desde un plano z_0 hasta un plano z_1 se describe como:

$$\begin{pmatrix} x' \\ \theta' \end{pmatrix} = T \begin{pmatrix} x \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix} = \begin{pmatrix} x - t\theta \\ \theta \end{pmatrix}. \quad (2.4)$$

Se puede apreciar que, en este caso, los rayos no alteran su dirección cuando se propagan libremente pero sí cambian su posición, tal y como se puede comprobar en la Figura 2.5 (a).

La matriz que describe la alteración que sufre un rayo al atravesar una lente delgada se expresa como:

$$L_f = \begin{pmatrix} 1 & 0 \\ 1/f & 1 \end{pmatrix}, \quad (2.5)$$

siendo f la focal de la lente. Por tanto, la transformación sufrida por los rayos al atravesar una lente delgada se describe como:

$$\begin{pmatrix} x' \\ \theta' \end{pmatrix} = L_f \begin{pmatrix} x \\ \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1/f & 1 \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ x/f + \theta \end{pmatrix}. \quad (2.6)$$

En este caso se observa que los rayos que atraviesan la lente sufren un cambio en su dirección, pero no en su posición, tal y como se puede comprobar en la Figura 2.5 (b).

Estas dos transformaciones son suficientes para describir los cambios que pueden llegar a sufrir los rayos de luz, procedentes de una escena 3D, en su interacción con sistemas de InI, y otras variantes de estos sistemas.

Captura de la función plenóptica

Para poder replicar la función plenóptica de una escena 3D, en primer lugar, es necesario contar con algún método que permita obtener y capturar información sobre la misma. A pesar de que en la actualidad se dispone de muchos dispositivos capaces de detectar y registrar la luz, obtener información sobre la función plenóptica no es una tarea sencilla. En esta Tesis, nos centraremos en dos técnicas que permiten obtener el flujo de luz procedente de una escena 3D: la InI y la fotografía plenóptica.

A continuación, explicaremos las dos técnicas de captura mencionadas. No obstante, para una mejor comprensión de ambas, es conveniente analizar en primer lugar cómo captura la función plenóptica la cámara convencional. Este estudio expondrá las complicaciones a las que se enfrentan los dispositivos que buscan capturar la información sobre el flujo de luz, a la vez que facilitará la comprensión de las técnicas de captura de InI y fotografía plenóptica.

3.1 Cámara convencional

Para este análisis se considerará una versión simplificada, pero sin pérdida de generalidad, de la estructura de una cámara fotográfica convencional. En general, una cámara se puede representar esquemáticamente mediante un sensor, que registra la luz recibida, y una lente convergente¹,

¹Esta lente representa al objetivo de la cámara que, en realidad, está compuesto por un conjunto de lentes.

la lente de la cámara (LC), que desvía la luz para formar imágenes en el sensor, ver Figura 3.1.

La LC proporciona una imagen de cada plano de la escena 3D. La posición de cada plano imagen, a' , puede ser determinada a partir de la posición de su correspondiente plano objeto, a , mediante la relación de conjugación de Gauss:

$$-\frac{1}{a} + \frac{1}{a'} = \frac{1}{f}, \quad (3.1)$$

donde f es la distancia focal de la LC. El plano de la escena 3D conjugado con el plano del sensor se denomina *plano de referencia* (PR). Los rayos procedentes de los puntos de la escena 3D situados en el PR convergerán en el plano del sensor, formando una imagen enfocada. Sin embargo, los rayos procedentes de puntos pertenecientes al resto de planos de la escena 3D no convergen en el plano del sensor y, por tanto, producen una imagen desenfocada en el mismo, ver los rayos azules en la Figura 3.1.

La función del sensor es registrar la luz que ha atravesado la LC. En esta Tesis se considerará que los sensores son de tipo CCD o CMOS, y que, por tanto, están compuestos por células fotoeléctricas o píxeles, de tamaño δ . Debido al tamaño finito de los píxeles, la imagen desenfocada de puntos pertenecientes a planos alejados del PR provocará la misma respuesta en el sensor que puntos situados en el PR si dicha imagen es menor que el tamaño δ de los píxeles del sensor. En consecuencia, se define un intervalo axial en el espacio objeto, denominado profundidad de campo Δ_c , que produce sobre el sensor imágenes igual de nítidas que los puntos del plano enfocado y puede expresarse como [41]:

$$\Delta_c = \frac{2 \phi a \delta |M|}{\delta^2 - |M|^2 \phi^2}, \quad (3.2)$$

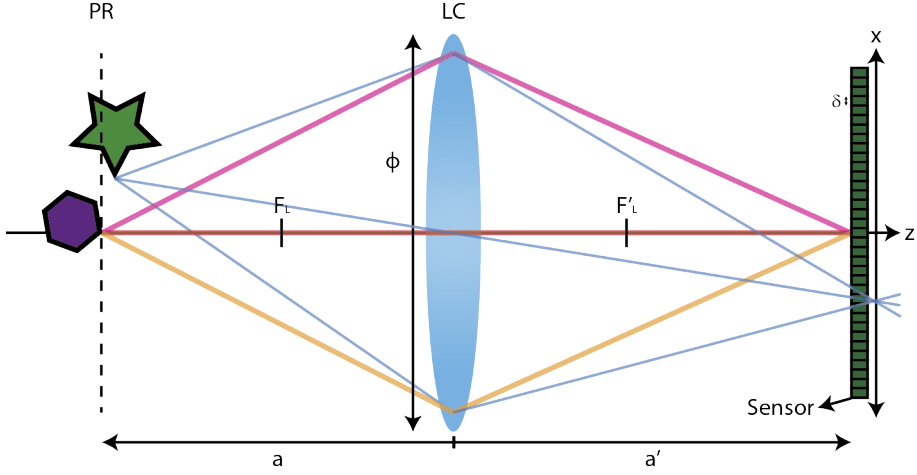


Figura 3.1.: Esquema de una cámara fotográfica convencional. F_L y F'_L se corresponden con los planos focal objeto y focal imagen de la LC, respectivamente; ϕ representa la apertura de la lente mientras que δ se corresponde con el tamaño del pixel del sensor de la cámara.

donde $M = a'/a$ es el aumento lateral entre el PR y el plano del sensor y ϕ representa la apertura de la lente.

Los píxeles del sensor integran el valor de la radiancia de todos los rayos que inciden sobre ellos para formar una imagen 2D, independientemente de su dirección de incidencia. Esto último tiene una consecuencia importante: la información direccional del flujo de luz es destruida cuando es registrado con una cámara convencional. La irradiancia de la imagen generada por este proceso se puede describir matemáticamente mediante la ecuación:

$$I(x, y) = \iint L(x, y, \theta, \varphi) d\theta d\varphi. \quad (3.3)$$

En la Figura 3.2 se puede apreciar la representación del proceso de formación de una imagen en el diagrama espacio-angular. En la Figura 3.2 (a) se muestran, en el plano del sensor de la cámara, los rayos resaltados

en la Figura 3.1. En las Figuras 3.2 (b) y (c) se proyectan todos los rayos sobre el eje x , representando el proceso de integración que realizan los píxeles para formar la imagen. Si nos fijamos en los rayos amarillo, rosa y rojo, se puede ver claramente como la información individual que transportaban se pierde cuando el pixel central suma sus radiancias.

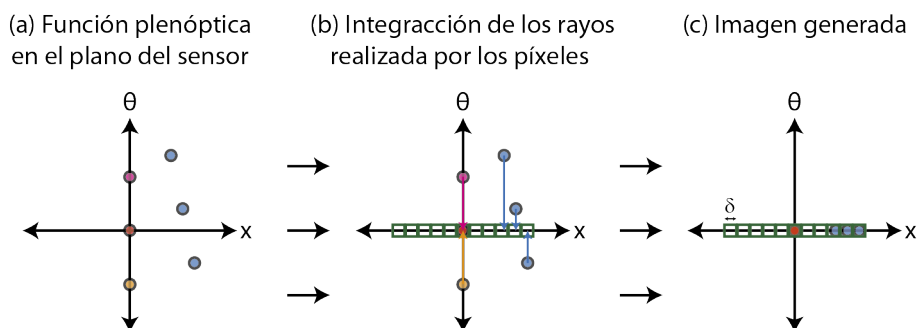


Figura 3.2.: (a) Representación espacio-angular de la función plenóptica capturada por la cámara de la Figura 3.1 en el plano del sensor de la cámara. (b) y (c) proceso de generación de la imagen 2D.

La cámara convencional es un dispositivo ampliamente utilizado para registrar la luz que proviene de una escena 3D. Aunque cumple con su objetivo de obtener una fotografía o perspectiva de la escena, no tiene la capacidad, por sí misma, de capturar la función plenóptica. Por tanto, es necesario utilizar otras técnicas o dispositivos más complejos que no sólo registren la irradiancia, sino también el mapa de radiancias, es decir, la energía que viaja a través de cada rayo o conjunto de rayos.

3.2 InI

Una de las técnicas con la capacidad de obtener información sobre la distribución del flujo de la luz es la InI, basada en la fotografía integral propuesta por Lippmann en 1908 [11]. El método de captura de esta técnica consiste en obtener distintas perspectivas de la escena 3D mediante una *matriz de lentes* (ML) y un sensor. A continuación, explicaremos las características y la configuración básica de un sistema de captura de InI.

Para facilitar la comprensión de sus características, intentaremos que el esquema sea lo más próximo posible al caso de la cámara convencional.

Consideremos la configuración de la cámara convencional representada en la Figura 3.1. Para implementar un sistema de captura de InI comparable a la cámara convencional, se reemplaza la LC por una ML. Las lentes de la matriz son de menor tamaño, p , y focal, f , que la lente a la que sustituyen pero, en su conjunto, ocupan la misma área que ésta. Adicionalmente, para mantener la equivalencia con el sistema de la Figura 3.1, el sensor de la cámara se desplaza para que el plano de la escena 3D conjugado con él, el PR, sea el mismo que en el caso de la cámara convencional.

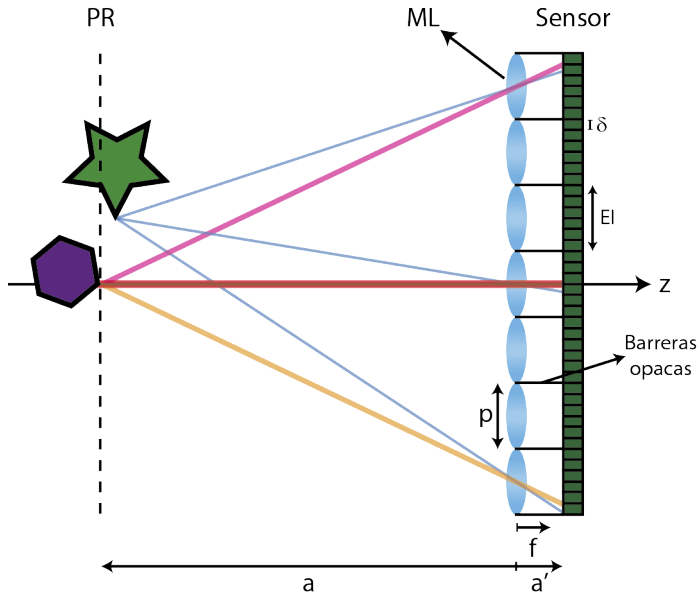


Figura 3.3.: Configuración básica del proceso de captura de un sistema de InI.

Los cambios que hemos descrito transforman la cámara convencional de la Figura 3.1 en el sistema de captura de InI mostrado en la Figura 3.3. En este nuevo sistema de captura, cada lente de la matriz forma una perspectiva distinta de la escena 3D sobre los píxeles del sensor. En este caso y de acuerdo con la ecuación (3.2), la profundidad de campo de las

perspectivas capturadas será mayor que la profundidad de campo de la cámara convencional. A las perspectivas individuales de la escena 3D se las conoce como imágenes elementales (EI, del inglés *elemental images*) y al conjunto de todas ellas se le denomina imagen integral (del inglés, *integral image*). Para evitar el solapamiento entre las distintas imágenes registradas, se sitúan barreras opacas que dividen el sensor en distintas celdas individuales de tamaño igual al de las lentes de la matriz, ver Figura 3.3. En las Figura 3.4 (a) y (b) mostramos, respectivamente, la fotografía convencional y la imagen integral de una misma escena 3D.

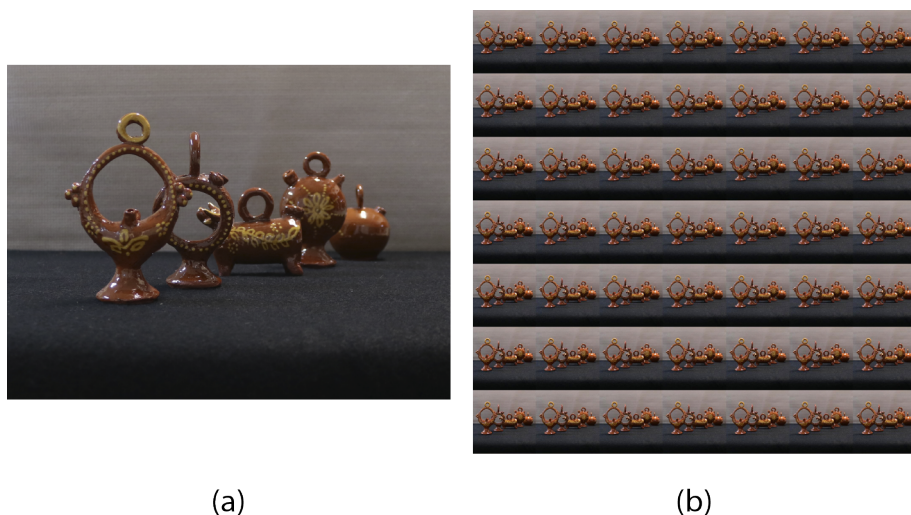


Figura 3.4.: (a) Fotografía convencional de una escena 3D. (b) Imagen integral de la escena mostrada en (a). Esta imagen integral está compuesta por 7×7 EIs.

Como se puede apreciar si comparamos las Figuras 3.1 y 3.3, en un sistema de captura de InI se ha producido una clasificación direccional del flujo de luz en los píxeles del sensor. Fijémonos en los rayos rosa, rojo y amarillo que se muestran en las dos figuras. En ambos casos, estos rayos impactan a la misma altura del plano en el que se encuentra la lente o la ML. En la cámara convencional, Figura 3.1, los tres rayos alcanzan el mismo pixel tras refractarse, independientemente de su ángulo de incidencia con el plano de la lente. Por tanto, cuando el pixel suma el valor de sus radiancias resulta imposible recuperar la información

transportada por cada rayo. Sin embargo, en el sistema de InI, Figura 3.3, debido a la existencia de la ML, la altura a la que inciden los rayos en el plano de la matriz y, por tanto, su inclinación, sí que son relevantes. Los rayos con distintas inclinaciones inciden en lentes diferentes, que los dirigen a los píxeles situados detrás de ellas. De este modo, cada rayo impacta, en función de su inclinación, en un pixel distinto del sensor, lo que permite preservar su información individual. Por tanto, podemos afirmar que, efectivamente, un sistema de InI permite obtener información de la función plenóptica.

Pero, haber conseguido información direccional de la función plenóptica ha tenido un precio: ahora, cada perspectiva de la escena 3D sólo utiliza una fracción de los píxeles totales del sensor. Como el número total de píxeles es fijo, incrementar el número de vistas capturadas de la escena 3D (información angular) reducirá el número de píxeles de cada perspectiva (información espacial). Este compromiso entre la información espacial y angular registradas es característico de los sistemas de captura de InI y, como veremos más adelante, de la fotografía plenóptica, siendo el ratio dependiente del número de lentes y de píxeles tras cada lente. En el caso de InI, la información angular capturada depende del número de lentes de la matriz, mientras que la información espacial depende del número de píxeles tras cada lente. Asimismo, la resolución espacial y angular con la que se captura esta información depende, respectivamente, del tamaño de los píxeles del sensor y de la separación angular de las lentes con respecto al punto central del PR.

Antes de analizar cómo se muestrea la función plenóptica en un sistema de InI conviene mencionar algunas cuestiones adicionales. En general, en InI se considera que las lentes que conforman la ML son delgadas y, por tanto, se aproxima gran parte de su comportamiento al de un estenope². En concreto, se considera que una lente, a pesar de que recoge luz en toda su superficie, sólo acepta haces de luz cuyo rayo central atraviesa el

²Pequeño agujero circular en la superficie de un material opaco que permite el paso de luz.

centro de la lente y, por tanto, no es desviado. Adicionalmente, se asume que no hay espacios vacíos entre las lentes de la matriz y que, por tanto, la separación entre lentes es igual a su tamaño, p . En la Figura 3.5 se muestra el esquema de una configuración de captura de InI que sigue este criterio.

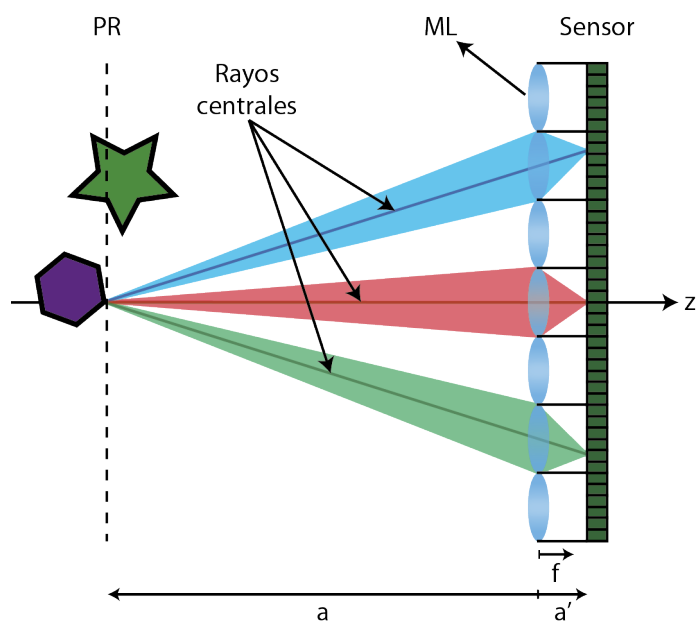


Figura 3.5.: Configuración básica de un sistema de InI bajo la aproximación de la lente como un estenopeo.

Para finalizar este punto, cabe mencionar que la configuración de captura explicada en este capítulo no es la única. Existen otras configuraciones de captura de InI que no requieren del uso de matrices de lentes. Por ejemplo, se pueden utilizar estenopes [50], que son más económicos que las matrices de lentes. Sin embargo, los estenopes reducen considerablemente la cantidad de luz que se registra en el sensor. Otra opción consiste en reemplazar las lentes y el sensor por cámaras fotográficas. En este caso, la resolución espacial de las perspectivas suele ser grande, ya que se usa la resolución completa del sensor en cada una de ellas. Además, son dispositivos fáciles de obtener y manipular, siendo posible cambiar con facilidad el PR o la focal del sistema de lentes. Sin embargo, la principal

desventaja de esta opción es lo costoso y aparatoso que puede ser una matriz de cámaras. Afortunadamente, en lugar de una matriz se puede usar una única cámara que se desplaza, mediante el uso de una guía o motores, para capturar las distintas perspectivas de la escena 3D. Este último método se conoce como captura de InI de apertura sintética [51], [52] y su principal desventaja es que imposibilita realizar una captura en tiempo real.

3.2.1 Muestreo de la función plenóptica en un sistema de InI

A continuación, se analizará el muestreo de la función plenóptica realizado por un sistema de captura de InI. En la Figura 3.6 se muestra el esquema de un sistema de captura de InI (a) y la representación espacio-angular de la función plenóptica muestreada por dicho sistema (b). Como se aprecia en la Figura 3.6 (b), podemos identificar la información contenida tras una lente, la EI, como una columna de puntos en el diagrama espacio-angular. Cabe señalar que para no complicar en exceso la representación, no se ha tenido en cuenta el tamaño del pixel, δ .

Como sólo se consideran los rayos que atraviesan el centro de las lentes, estos serán los únicos (en nuestra aproximación) registrados por el sensor, ver Figura 3.6 (a). Por esta razón, un sistema de captura de InI muestrea la función plenóptica en el plano de la ML. El muestreo espacial de la función plenóptica depende del número de lentes de la matriz, siendo la resolución de dicho muestreo igual a la separación p entre los centros de las lentes. Por otro lado, el muestreo angular de la función plenóptica es igual al número de píxeles tras cada lente, pues estos limitan el número de haces con distinta inclinación que pueden ser registrados, siendo la resolución angular con la que se realiza este muestreo igual al cociente entre el tamaño del pixel, δ , y la separación entre los píxeles y la ML, a' .

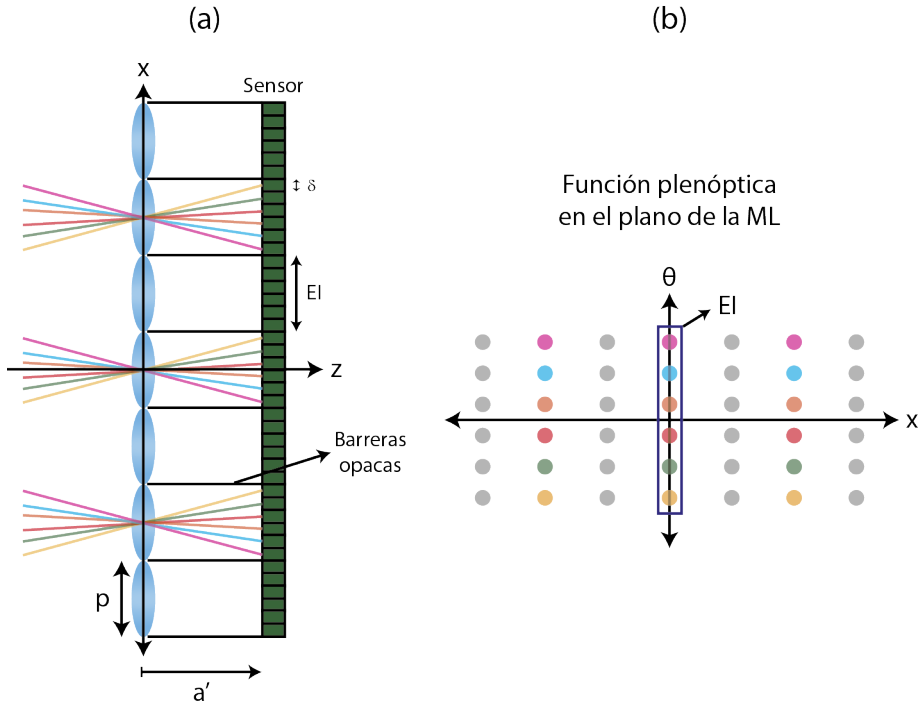


Figura 3.6.: (a) Muestreo de la función plenótica realizado por un sistema de InI. (b) Diagrama espacio-angular de la función plenótica muestreada en (a). Cada columna de puntos se corresponde con los rayos que componen una EI.

Al igual que sucedía con la información de la escena 3D capturada, existe un compromiso entre la información espacial y angular de la función plenótica que se registra en el sensor, dependiente también del número de lentes y de píxeles tras cada lente.

3.3 Fotografía plenótica

El segundo método que permite obtener información sobre la distribución del flujo de luz que se analizará en esta Tesis es la fotografía plenótica [23], [46], [47]. Este método está basado en la propuesta que realizaron Adelson y Wang en 1992 para capturar la función plenótica

[21]. El método de captura de esta técnica consiste en distribuir los rayos de luz, procedentes de la escena 3D, en función de su inclinación mediante la inserción de una ML entre el sensor y la lente de una cámara convencional. A continuación, explicaremos las características y la configuración básica de este rediseño de la cámara convencional, conocido como cámara plenóptica 1.0.

Consideremos de nuevo la configuración de captura de la cámara convencional representada en la Figura 3.1, para implementar un sistema de captura de fotografía plenóptica comparable a la cámara convencional, se inserta una ML, de tamaño p y focal f , en el plano del sensor mientras que éste se desplaza una distancia g , ver Figura 3.7.

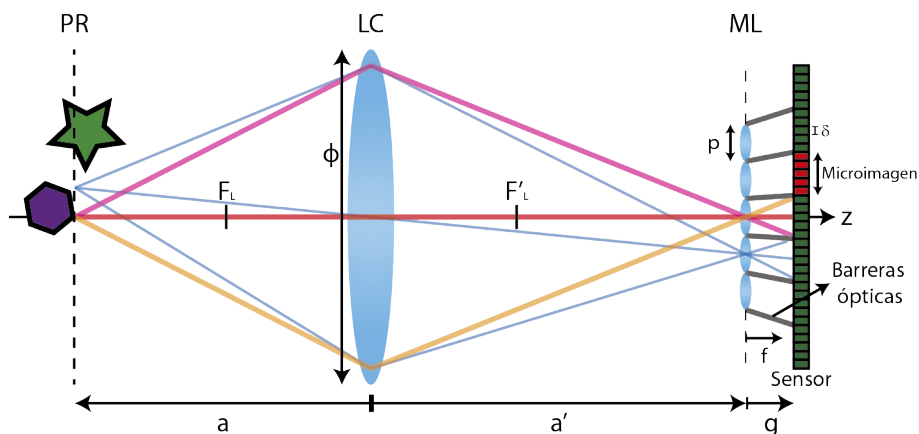


Figura 3.7.: Configuración básica de una cámara plenóptica.

Para el correcto funcionamiento de este dispositivo, se deben satisfacer dos relaciones de conjugación. Por un lado, el plano de la ML debe encontrarse conjugado con un plano de la escena 3D a través de la LC. Como la ML ha reemplazado al sensor de la cámara convencional, este plano será el mismo que se encontraba conjugado con el sensor en la cámara convencional. Por motivos de conveniencia, y ya que en esta configuración la ML realiza un papel parecido al del sensor, también denominaremos a este plano, *plano de referencia* (PR). Por otro lado, el plano del sensor, g , tiene que encontrarse conjugado con el plano de la

LC a través de la ML. Como la distancia entre la ML y la LC, a' , suele ser muy grande en comparación con la focal de las lentes de la matriz, f , y teniendo en cuenta la siguiente relación de conjugación:

$$-\frac{1}{a'} + \frac{1}{g} = \frac{1}{f}, \quad (3.4)$$

la posición del sensor se puede aproximar por $g = f$ [53].

Una cámara plenóptica es un sistema de funcionamiento más complejo que una cámara convencional. Por una parte, la función de la LC de la cámara plenóptica es producir una imagen de la escena 3D en los alrededores de la ML [54]. Específicamente, se forma una imagen nítida del PR sobre el plano de la ML, mientras que las imágenes del resto de planos de la escena 3D se generarán en los alrededores de la matriz. Esto implica que, a diferencia de InI, cada lente de la matriz sólo recoge luz de una pequeña porción de la escena 3D. Por motivos de claridad, de aquí en adelante denominaremos *imagen flotante* a la imagen de la escena 3D proporcionada por la LC, ver Figura 3.8.

Por otra parte, como el plano del sensor se encuentra conjugado con la LC, y no con la escena 3D, cada lente de la matriz produce sobre los píxeles del sensor una imagen nítida de la apertura de la LC y una imagen totalmente desenfocada de la escena 3D. A las imágenes tras las lentes de la matriz se les denomina microimágenes, y al conjunto de todas ellas imagen plenóptica, ver Figura 3.9. Como se puede ver en la figura, el aspecto de una imagen plenóptica se corresponde con el de una imagen desenfocada de la escena 3D que se encuentra dividida en sectores por imágenes nítidas de la apertura, las microimágenes. Conviene mencionar que, al igual que en InI, se consideran solo los haces de luz cuyo rayo central atraviesa el centro de las lentes. Asimismo, se asume que no hay espacios vacíos entre las lentes de la matriz y que, por tanto, la distancia entre centros es igual a su tamaño, p .

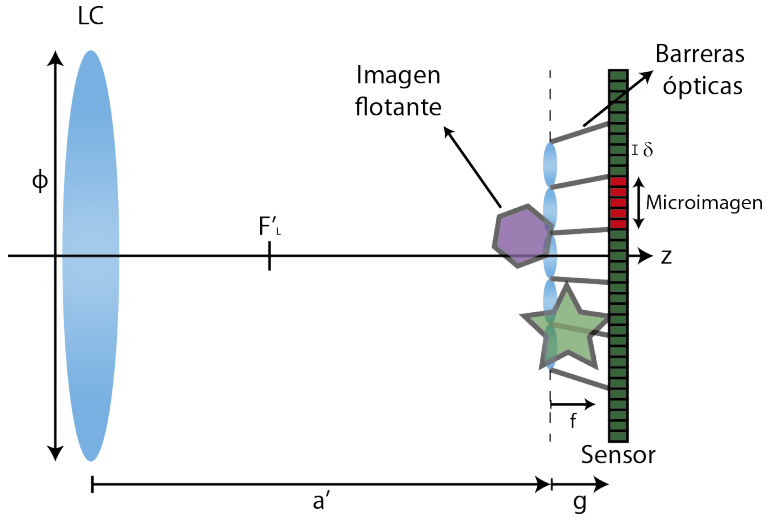


Figura 3.8.: Ejemplo de la imagen 3D que se forma en los alrededores de la ML en una captura plenótica, aquí denominada imagen flotante.

A diferencia de la técnica de InI, en una cámara plenótica no se sitúan barreras opacas para separar las distintas imágenes. En su lugar, para evitar el solapamiento entre las distintas microimágenes, se controla la apertura, ϕ , de la LC mediante un diafragma. Este proceso de ajuste genera, gracias al viñeteado, un conjunto de barreras ópticas oblicuas que separa las microimágenes, ver Figura 3.7. Para que las microimágenes sean tangentes entre sí y se aproveche el máximo número de píxeles del sensor, ver Figura 3.10, el tamaño del diafragma debe satisfacer la siguiente relación entre la apertura de la LC y las lentes de la matriz:

$$\frac{\phi}{a'} = \frac{p}{f}. \quad (3.5)$$

Hasta ahora hemos descrito la configuración básica de la cámara plenótica 1.0 y algunas de sus peculiaridades. Si, al igual que en el caso de InI, comparamos la configuración básica de esta técnica, Figura 3.7, con la cámara convencional, Figura 3.1, vemos que también se produce una clasificación direccional del flujo de luz en los píxeles del sensor.

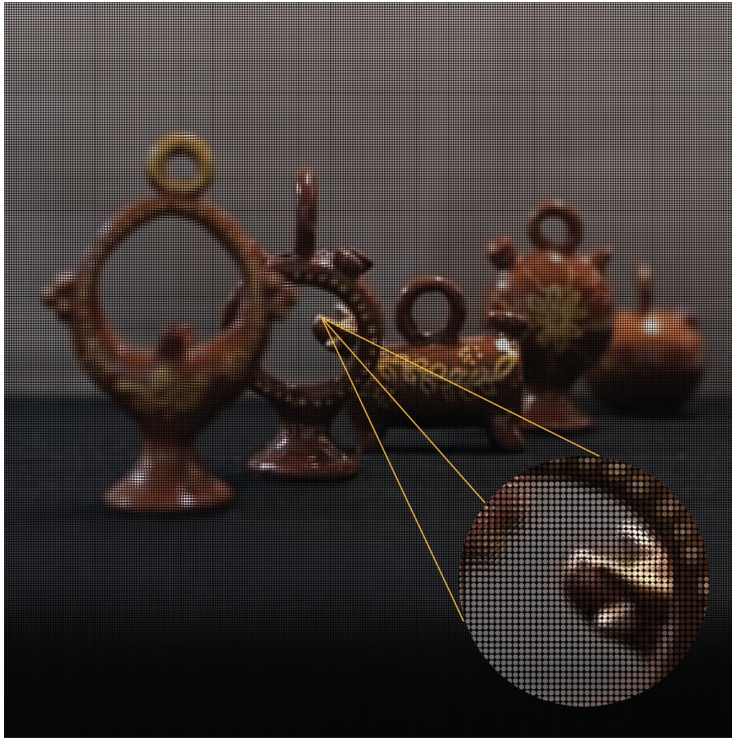


Figura 3.9.: Imagen plenótica compuesta por 300×300 microimágenes. El PR de esta imagen se encuentra sobre el tercer botijo.

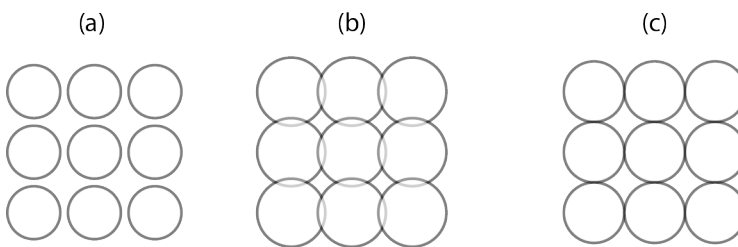


Figura 3.10.: Microimágenes de una captura plenótica. Dependiendo de la relación entre las aperturas relativas de la LC y las lentes de la matriz puede suceder que: (a) se desperdicien píxeles del sensor, (b) se produzca solapamiento entre microimágenes, (c) se obtengan microimágenes tangentes, aprovechando al máximo los píxeles del sensor.

Fijémonos en los rayos rosa, rojo y amarillo de la Figura 3.7, los tres originados en el PR. Mientras que en la cámara convencional impactaban en un mismo pixel del sensor, ahora atraviesan el centro de una lente de la matriz, alcanzando píxeles distintos de una microimagen. Por tanto, con esta técnica también se preserva la información individual de los rayos, ya que cada uno impacta en un pixel distinto del sensor en función de su inclinación. No obstante, como se puede apreciar al comparar las Figuras 3.5 y 3.7, la distribución del flujo de luz en el sensor de una cámara plenótica es distinta a la de un sistema de InI. Mientras que en una cámara plenótica los rayos originados en un punto situado en el PR se registran en la misma microimagen, en el caso de InI se registran en varias EIs. Hay que destacar que esta diferencia en la distribución disminuye para planos alejados del PR. En este caso, la imagen de estos planos ya no se forma sobre la ML de la cámara plenótica y, por tanto, los rayos procedentes de ellos se registrarán en más de una microimagen, ver la Figura 3.11.

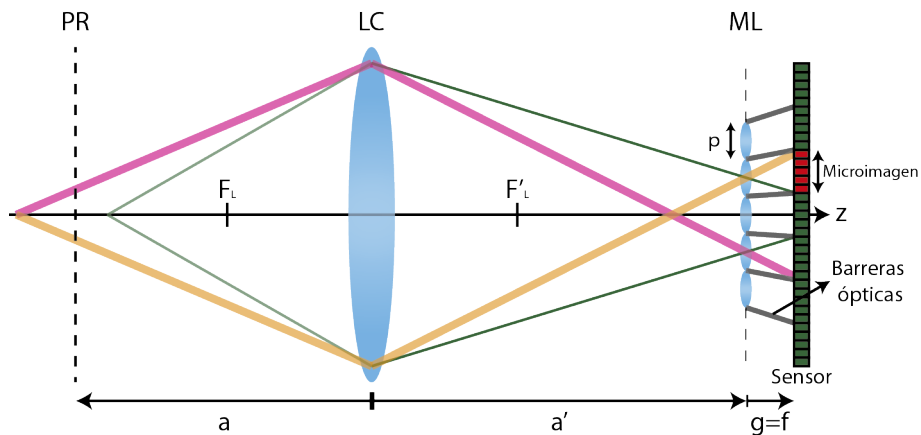


Figura 3.11.: Esquema de la captura plenótica de planos de la escena 3D alejados del PR.

Al igual que en InI, en una cámara plenótica también existe un compromiso entre la cantidad de información espacial y angular que se puede obtener de la escena 3D debido al número finito de píxeles disponibles. Pero, en este caso, y a diferencia de InI, son las lentes de la matriz las que realizan un muestreo espacial de la escena 3D, de la imagen flotante

para ser más exactos. Por otro lado, los píxeles del sensor situados tras las lentes registran rayos con distinta inclinación de los puntos de la imagen flotante muestreados, es decir, registran las vistas. En consecuencia, el ratio de captura de la información espacial y angular de la escena 3D es transpuesto al de la técnica de InI [30], [54]. Asimismo, la resolución espacial y angular con la que se captura esta información depende, respectivamente, del tamaño de las lentes y de la separación angular de los píxeles de cada microimagen con respecto al centro de su correspondiente lente.

Antes de analizar cómo se muestrea la función plenóptica en un sistema de captura plenóptico conviene mencionar algunas cuestiones adicionales. Las lentes de la matriz se encuentran conjugadas con la escena 3D y, como hemos visto, son las encargadas de muestrear la información espacial de la misma, como los píxeles de una cámara convencional. Por tanto, la resolución y la profundidad de campo del sistema plenóptico dependen de las lentes de la matriz, afectando al aspecto visual de la imagen plenóptica. Por ejemplo, la imagen del PR aparecerá enfocada en la imagen plenóptica, mientras que la imagen de los planos de la escena 3D que no se encuentran conjugados con la ML se verá desenfocada, ya que rayos que proceden de un mismo punto de la escena 3D impactarán en más de una lente, ver Figura 3.11. Teniendo esto en cuenta, se podría mejorar la profundidad de campo de la imagen plenóptica cerrando el diafragma de la LC para obtener microimágenes como en la Figura 3.10 (a). En este caso, se recortarían rayos procedentes de puntos de la imagen flotante que impactarían en más de una lente, aunque se registrarían menos vistas de la escena 3D.

Otra cuestión que conviene mencionar es que las vistas capturadas por un sistema de fotografía plenóptica, denominadas subimágenes [23], [54], se pueden generar de manera computacional. Para obtener un subimagen, se construye una imagen 2D extrayendo los píxeles que ocupan la misma posición relativa en cada microimagen de la imagen plenóptica, tal y

como se muestra en la Figura 3.12. Como cada pixel de una microimagen almacena información angular distinta de la escena, cambiar el pixel que se extrae de las microimágenes cambia la subimagen o vista generada. El número de subimágenes o perspectivas que se pueden obtener de una imagen plenóptica depende del número de píxeles por microimagen. Por otro lado, el número de píxeles que contiene cada subimagen depende de la cantidad total de microimágenes de la imagen plenóptica. En la Figura 3.13 se muestran las subimágenes obtenidas a partir de la imagen plenóptica de la Figura 3.9.

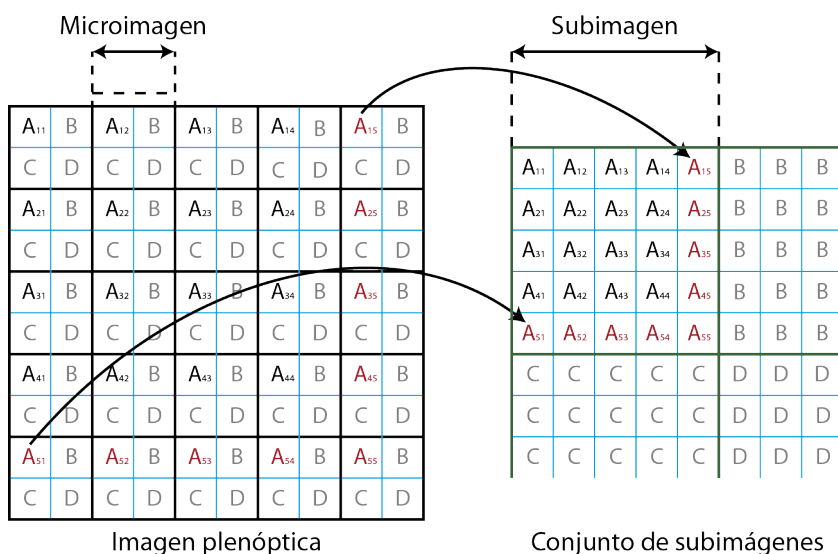


Figura 3.12.: Proceso de generación de las subimágenes.

Para finalizar este punto, cabe mencionar que existe otra configuración de captura similar a la cámara plenóptica 1.0, conocida como cámara plenóptica 2.0 [46], [55]. En este diseño alternativo de la cámara plenóptica, las relaciones de conjugación son diferentes al caso 1.0. En la versión 2.0, la imagen de la escena 3D se forma entre la LC y la ML, mientras que el sensor se conjuga con un plano de dicha imagen. El resultado final es una configuración de captura muy parecida a la InI que obtiene imágenes con viñeteado. En este caso, el ratio entre la información espacial y angu-



Figura 3.13.: Subimágenes de la imagen plenótica de la Figura 3.9.

lar capturada de la escena 3D se encuentra a medio camino entre InI y plenótica 1.0.

3.3.1 Muestreo de la función plenótica en un sistema de fotografía plenótica

A continuación, se analizará el muestreo de la función plenótica realizado por un sistema de captura plenótico 1.0 que, como se verá en los siguientes párrafos, es muy similar al realizado por un sistema de InI. En la Figura 3.14 se muestra el esquema de captura de una cámara plenótica (a) y la representación espacio-angular de la función plenótica muestreada por dicho sistema (b). Como se aprecia en la Figura 3.14 (b), podemos identificar la información contenida tras una lente, la microimagen, como

una columna de puntos en el diagrama espacio-angular, mientras que cada fila representa una vista o subimagen diferente de la escena 3D. Cabe señalar que para no complicar en exceso la representación, no se ha tenido en cuenta el tamaño del pixel, δ .

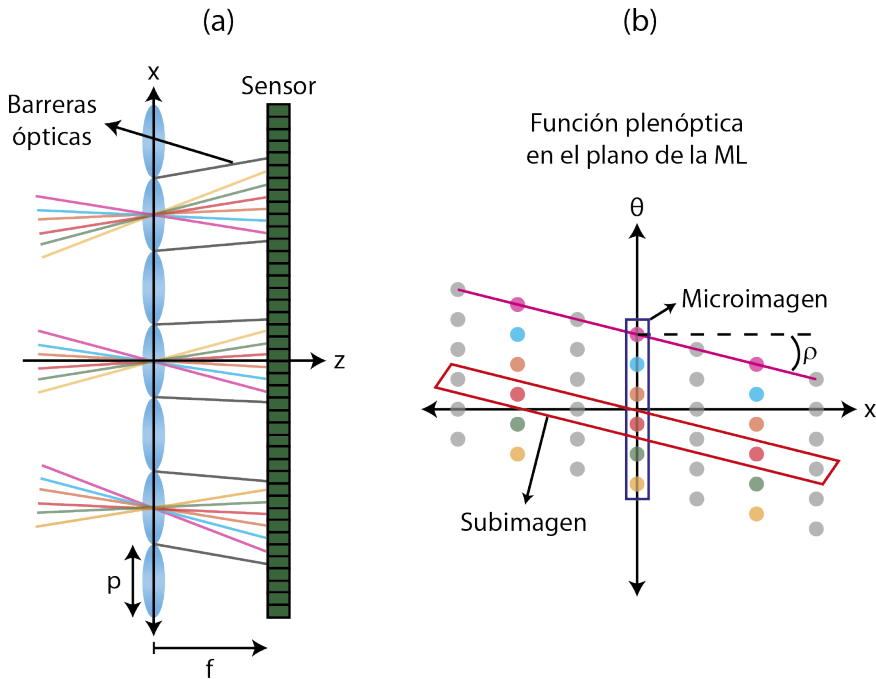


Figura 3.14.: (a) Muestreo de la función plenóptica por un sistema de captura plenóptico. (b) Diagrama espacio angular de la función plenóptica muestreada en (a). Cada columna de puntos se corresponde con los rayos que componen una microimagen.

Al igual que en InI, sólo se consideran los rayos que atraviesan el centro de las lentes y, por tanto, estos serán los únicos (en nuestra aproximación) registrados por el sensor, ver Figura 3.14 (a). Por esta razón, un sistema de captura plenóptico también muestrea la función plenóptica en el plano de la ML. En consecuencia, el muestreo espacial y angular de la función plenóptica realizado por una cámara plenóptica tiene la misma naturaleza que el realizado por un sistema de InI. En este caso, el muestreo espacial y angular también depende del número de lentes de la matriz y del número de píxeles tras cada lente, respectivamente. Asimismo, la resolución del

muestreo también es determinada por la separación p entre los centros de las lentes, para el caso del muestreo espacial, y el cociente entre el tamaño del pixel δ y la separación entre los píxeles y la ML, g , para la resolución angular. Sin embargo, a diferencia de la técnica de InI, existe una cizalladura en la función plenótica registrada. Esta cizalladura, de ángulo $\rho = -1/a'$, se produce debido a que cada lente de la matriz recibe rayos con inclinación ligeramente diferente, como se puede apreciar en la Figura 3.14 (a).

Al igual que sucedía con la información de la escena 3D capturada, existe un compromiso entre la información espacial y angular de la función plenótica que se registra en el sensor, dependiente también del número de lentes y de píxeles tras cada lente.

3.4 Equivalencia entre InI y fotografía plenótica

A pesar de que la InI y la fotografía plenótica difieren en su implementación, existe una relación de equivalencia entre ambos sistemas de captura. A continuación, describimos brevemente dicha relación.

Si se analiza la configuración básica de una cámara plenótica, Figuras 3.7 y 3.8, vemos que una cámara plenótica es, en realidad, un sistema de captura de InI que utiliza una lente para acercar la escena 3D a la ML³. En concreto, este sistema de InI se encuentra sumergido en la imagen de la escena 3D, con planos de esta última antes, sobre y tras la ML. Es evidente que las características de este sistema sumergido serán las mismas que las de una cámara plenótica. Por ejemplo, en el sistema de InI sumergido, las imágenes obtenidas tienen viñeteado, se encuentran separadas por barreras oblicuas y las lentes de la matriz sólo recogen luz

³De hecho, algunos autores denominan InI de campo lejano a la técnica de la fotografía plenótica [24], [26].

de una pequeña porción de la escena 3D. Si tenemos estas características en cuenta, sería posible realizar una captura de InI que obtuviera en su sensor algo parecido a una imagen plenóptica. Por ejemplo, se podría implementar un sistema de captura de InI con el sensor conjugado con el infinito y en el que la escena 3D se encontrara lo suficientemente cerca de la ML para que cada punto de la misma sólo se registrara en una o pocas imágenes. Hay que destacar que, en este caso, las barreras no serían oblicuas.

Otra forma de mostrar la relación de equivalencia entre la InI y la fotografía plenóptica aparece al analizar las subimágenes de una imagen plenóptica. Como se puede apreciar en la Figura 3.15 (a), los píxeles que componen una subimagen⁴ contienen información de todos los rayos que han atravesado la misma región o sub-apertura de la LC. Por tanto, se puede considerar que las subimágenes son perspectivas de la escena 3D que han sido capturadas por un sistema de InI situado en el plano de la LC [54]. El tamaño y la posición de las lentes de la matriz de este sistema de InI equivalente se corresponderían con el tamaño y la posición de las sub-aperturas de la LC. Conviene destacar que, a diferencia de un sistema de InI convencional, las barreras que separarían las lentes serían oblicuas [54], o lo que es lo mismo, que los ejes ópticos de las lentes empleadas en la captura se cruzarían en el PR, como se puede apreciar en la Figura 3.15 (b).

Concluimos este apartado con una aclaración. Como acabamos de comentar, existe una relación de equivalencia entre los sistemas de captura de InI y fotografía plenóptica. Para evitar posibles confusiones entre las imágenes que obtienen ambos sistemas de la escena 3D, de aquí en adelante haremos una distinción clara. Independientemente del sistema o método empleado para la captura, si las imágenes que se obtienen registran toda la escena 3D de interés, se denominarán EIs. Sin embargo, si las imágenes obtenidas solo contienen una pequeña parte de la escena,

⁴Recordemos que una subimagen se construye extrayendo los píxeles que ocupan la misma posición en las microimágenes.

o dicho de otra forma, si cada parte de la escena sólo es registrada en una o pocas imágenes, estas imágenes se denominarán microimágenes.

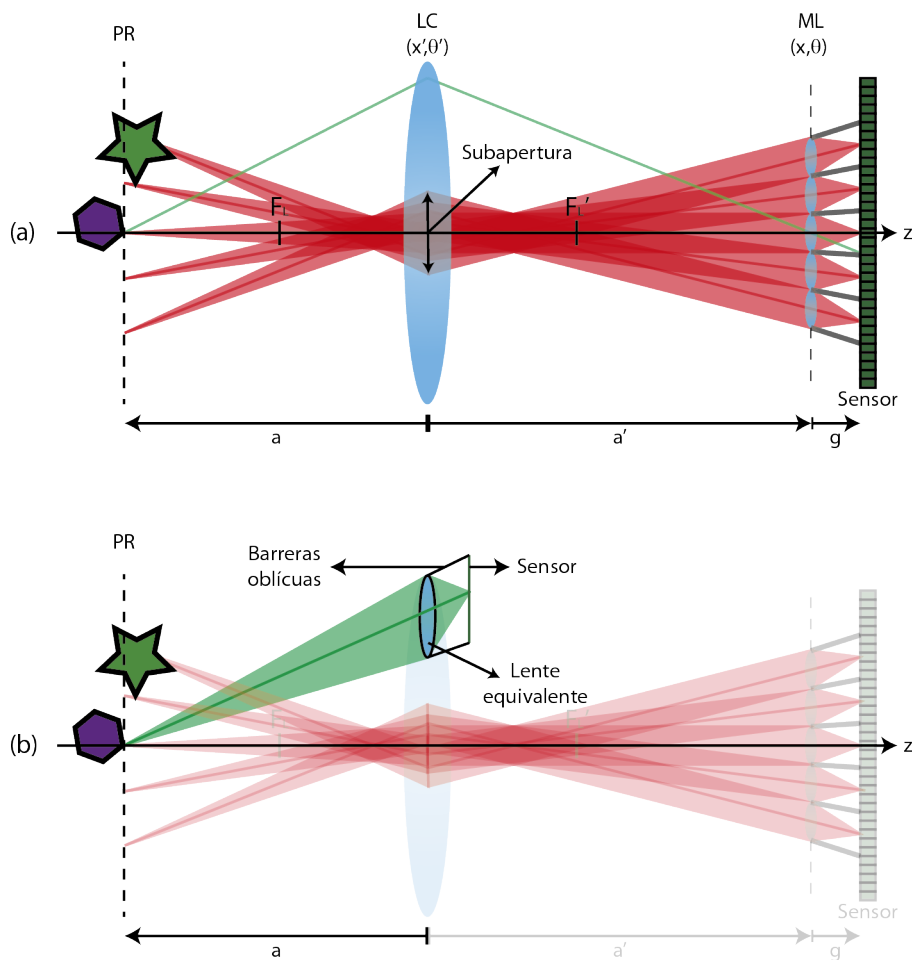


Figura 3.15.: (a) Sistema de captura plenóptico en el que se muestra que los píxeles en la misma posición relativa detrás de cada lente han atravesado la misma región de la LC. (b) Una lente del sistema de captura de InI equivalente que capturaría las subimágenes generadas por la cámara plenóptica.

3.5 Casos reales

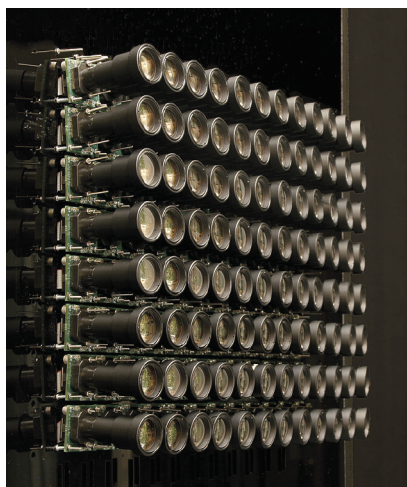
A lo largo de este capítulo se ha establecido un paralelismo entre los dos sistemas de captura de interés para esta Tesis, InI y cámara plenóptica, y la cámara convencional (Figura 3.1). El objetivo de este paralelismo ha sido facilitar la comprensión y análisis de los dos sistemas mencionados. Sin embargo, para completar la descripción, es necesario mencionar un aspecto no contemplado durante la explicación: las características de la ML (tamaño, focal y número de lentes) y del sensor empleados suelen variar considerablemente en función de si el sistema de captura implementado es un sistema de InI o una cámara plenóptica.

En la captura de InI, las matrices empleadas suelen tener un reducido número de lentes cuyas aperturas pueden variar dentro de un amplio rango, desde los milímetros hasta varios centímetros, lo que permite la obtención de un gran número de píxeles por cada EI. Por ejemplo, en el caso de realizar una captura de InI con una matriz de cámaras, ver Figura 3.16 (a)⁵, la cantidad de información y paralaje de la escena 3D capturada puede llegar a ser muy elevada. Destacar que, en este sistema, el tamaño total del sensor sería $N_H \times N_V$ veces más grande que el tamaño del sensor de una única cámara, donde N_H y N_V son el número de cámaras en horizontal y vertical utilizadas, respectivamente. Por otro lado, una cámara plenóptica es un sistema menos flexible, ya que el tamaño de la ML se encuentra fuertemente limitada por el tamaño y la estructura de la cámara en sí, ver Figura 3.16 (b)⁶. En consecuencia, en los sistemas plenópticos es habitual utilizar matrices de pequeñas dimensiones, con un tamaño similar al tamaño del sensor de la cámara empleada. Debido a ello, las lentes de la matriz son de apertura pequeña, suelen ser inferiores al milímetro, y las microimágenes que se obtienen tienen pocos píxeles.

⁵Imagen: <http://graphics.stanford.edu/projects/array/>.

⁶Imagen: <https://commons.wikimedia.org/w/index.php?curid=38422894>, By Morio - Own work, CC BY-SA 3.0.

Esta disonancia entre las características de los elementos comunes que se utilizan en las técnicas de InI y fotografía plenóptica afecta al tamaño físico del dispositivo de captura y a la cantidad de información almacenada, siendo ambos usualmente mayores en el caso de InI. A grandes rasgos, un sistema de InI permite obtener más paralaje y resolución de la escena 3D que una cámara plenóptica. Aun así, no se ha de menospreciar la principal ventaja de una cámara plenóptica: la capacidad de capturar la función plenóptica en un único disparo con un sistema compacto y portátil.



(a)



(b)

Figura 3.16.: (a) Sistema de captura de InI compuesto por una matriz de cámaras. (b) Cámara plenóptica comercial.

Reconstrucción de la función plenóptica

Obtener información sobre la función plenóptica de una escena 3D permite generar una imagen 3D de la misma. Para ello, se reconstruye el flujo de luz que emite la escena utilizando un monitor de InI. En este capítulo, explicaremos este proceso de reconstrucción. Además, también comentaremos las características de un monitor de InI y describiremos cómo un observador percibe la imagen 3D generada. Para finalizar, plantearemos las características que debe tener un monitor de InI para que las imágenes 3D que genere sean percibidas con buena calidad.

4.1 El monitor de InI

Cuando la técnica de InI fue originalmente propuesta por Lippmann [11], tenía un objetivo más ambicioso que el de meramente capturar el flujo de luz de una escena 3D: se planteó con la intención de desarrollar un método de generación de imágenes 3D. El método propuesto por Lippmann consistía en la reconstrucción del flujo de luz emitido por una escena 3D que había sido previamente capturado. El monitor de InI que se describe en este capítulo es una versión evolucionada del dispositivo empleado por Lippmann para reconstruir el flujo de luz.

La configuración básica de un monitor de InI se compone por un monitor (o pantalla pixelada) y una ML, que denominaremos de aquí en adelante

como *matriz de reconstrucción* (MR), situada delante de él, ver Figura 4.1.

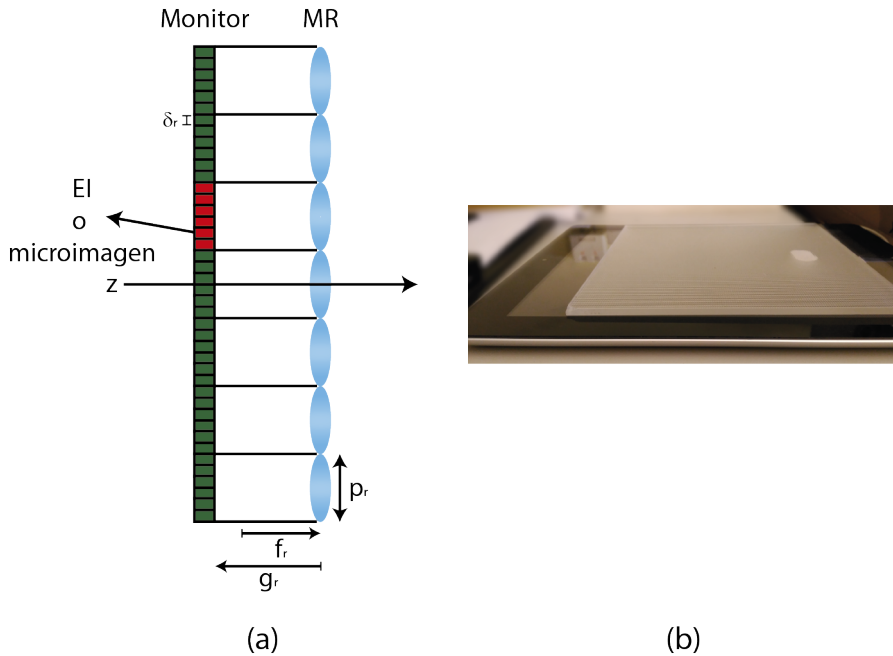


Figura 4.1.: (a) Configuración básica de un monitor de InI. f_r y p_r representan, respectivamente, la distancia focal y el tamaño de las lentes de la MR; g_r indica la distancia entre los píxeles del monitor y la MR; δ_r es el tamaño del píxel del monitor. (b) Fotografía de un monitor de InI compuesto por una tablet y una MR.

Para que un monitor de InI pueda reconstruir el flujo de luz emitido por una escena 3D, sobre los píxeles del monitor debe mostrarse la función plenóptica capturada de dicha escena, ya sea codificada en forma de EIs o en forma de microimágenes. A partir de la descripción anterior, y comparando la Figura 4.1 con los esquemas de los sistemas que vimos en las secciones 3.2 y 3.3, se deduce fácilmente que un monitor de InI es similar a un dispositivo de captura de la función plenóptica en el que se ha sustituido el sensor por una pantalla que emite luz. Por tanto, en lugar de registrar el flujo de luz en forma de EIs (o microimágenes), un monitor de InI emite el flujo de luz contenido en éstas. En la siguiente sección, veremos como la semejanza entre el monitor de InI y el sistema

que se empleó para capturar las imágenes que se muestran en él afecta al flujo de luz reconstruido.

4.1.1 Configuración simétrica y asimétrica de un monitor de InI

Si las imágenes, ya sean EIs o microimágenes, que se muestran en el monitor de InI se han obtenido con un sistema de captura en el que las características de la ML y el sensor son iguales a las empleadas en el monitor de InI, se dice que este último trabaja en configuración simétrica. Cuando la configuración de un monitor de InI es simétrica, éste es una réplica del sistema empleado en la captura. En consecuencia, los haces de luz emitidos por las imágenes formarán una réplica del flujo de luz que se obtuvo en el proceso de captura. Hay que destacar que si no hay barreras que separen las distintas imágenes, la luz emitida por estas últimas podría atravesar lentes no correspondientes. En este caso, se producirían flujos de luz secundarios que no serían réplicas exactas del flujo capturado. Trataremos este asunto con más detalle en la sección 4.2.1, de momento, nos centraremos en el flujo de luz que atraviesa las lentes adecuadas.

Que el flujo de luz emitido por un monitor de InI sea una réplica del capturado implica que, por el principio de reversibilidad de los rayos, los haces de luz se cruzarán en los puntos de la escena 3D que los habían originado en el proceso de captura. En las zonas donde se produce la intersección de estos haces de luz aumenta la densidad local de energía y, por tanto, se forma un punto luminoso que, a todos los efectos, se comporta como una fuente de luz. El conjunto de los puntos luminosos generados constituye una réplica auto-luminosa de la escena 3D original, es decir, una imagen 3D. Hay que tener en cuenta que en el caso en que se muestren en el monitor de InI microimágenes obtenidas con una cámara plenóptica, los rayos se cruzarán en los puntos donde se encontraba la imagen flotante, ya que es el flujo de luz de ésta, y no el de la escena 3D

original, el que se ha registrado en la imagen plenóptica. Esto implica que, en este caso, la imagen 3D se encontrará en los alrededores de la MR. Nótese que el plano de la imagen 3D que se forma sobre dicha matriz dependerá del PR de la cámara plenóptica que capturó la imagen. Por motivos de conveniencia, usaremos el término *PR de la imagen plenóptica* para referirnos al plano de la imagen 3D que se forma sobre la MR de un monitor de InI.

En la Figura 4.2 se muestra la captura del flujo de luz de una escena mediante un dispositivo de InI y su posterior reconstrucción gracias a un monitor de InI operando en configuración simétrica y con barreras separando las distintas imágenes. Del mismo modo, en la Figura 4.3 se muestra la captura del flujo de luz de una escena usando un dispositivo de fotografía plenóptica y su posterior reconstrucción mediante un monitor de InI operando en configuración simétrica. Conviene matizar que la formación de la imagen 3D es un proceso completamente independiente a la capacidad de las lentes de la MR de formar imágenes. Tal y como se puede apreciar en las Figuras 4.2 (b) y 4.3 (b), a pesar de que los haces de luz emitidos por el monitor de InI se cruzan en distintos planos para formar la imagen 3D, estos convergen en el plano conjugado con los píxeles a través de la MR, que denominamos plano de referencia imagen (PRI), z_r . La posición de dicho plano depende de la focal de las lentes, f_r , y de la distancia entre los píxeles del monitor y la MR, g_r , de la siguiente forma:

$$-\frac{1}{g_r} + \frac{1}{z_r} = \frac{1}{f_r}. \quad (4.1)$$

Como veremos más adelante, esta capacidad de producir imágenes provocará problemas en la visualización de la imagen 3D.

Desafortunadamente, en la práctica, es difícil implementar un monitor de InI que trabaje en configuración simétrica. En general, las matrices de

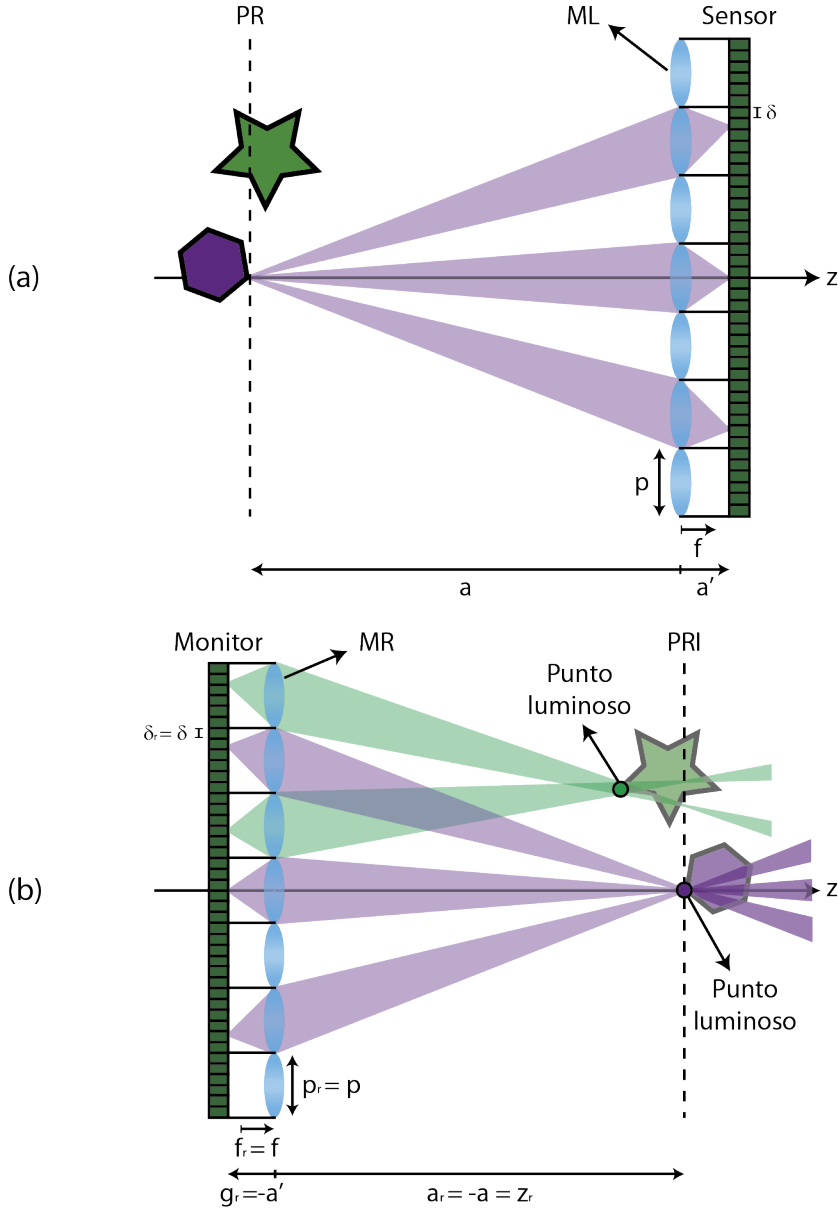


Figura 4.2.: (a) Configuración de captura de InI que emplea una ML y un sensor. (b) Reconstrucción del flujo de luz capturado en (a) mediante un monitor de InI en configuración simétrica. En este caso, los parámetros del monitor de InI, g_r , f_r , p_r y δ_r , son idénticos a sus homólogos del sistema de captura, g , f , p y δ . a_r indica la distancia a la que se reconstruye la imagen 3D de la MR.

lentes y los píxeles del monitor empleados en un monitor de InI suelen ser bastante diferentes a sus equivalentes utilizados en la captura de la función plenóptica. Por ejemplo, los dispositivos más comúnmente utilizados como monitores, como las pantallas LCD o LED de móviles, ordenadores y tabletas, tienen un número, tamaño y distribución de píxeles distinto a los píxeles presentes en el tipo de sensor que se emplea en la captura de la función plenóptica, como, por ejemplo, el sensor CCD o CMOS de una cámara. Del mismo modo, las MLs empleadas en el proceso de captura de la función plenóptica suelen tener parámetros diferentes a las MRs. Por ejemplo, las MRs tienen un tamaño y distancia focal mucho mayor que las matrices empleadas en el interior de una cámara plenóptica ya que, estas últimas, se encuentran fuertemente limitadas por el tamaño de la cámara. En consecuencia, no es raro que los monitores de InI utilicen elementos de características diferentes a los elementos usados en el proceso de captura de la función plenóptica. Cuando un monitor de InI trabaja en estas condiciones, que identificaremos como configuración asimétrica, el flujo de luz que reconstruye ya no es una réplica exacta del capturado. Las características del flujo de luz reconstruido en estas condiciones dependen de las diferencias existentes entre el sistema utilizado para la captura de las imágenes y el propio monitor de InI.

Si las diferencias son debidas a que el número y/o la focal de las lentes empleadas en el sistema de captura es diferente al número y/o focal de las lentes empleadas en el monitor de InI, se produce un cambio en el número y/o en la convergencia de los haces de luz emitidos con respecto a los capturados. No obstante, estas alteraciones no modifican el lugar dónde se cruzan los haces de luz y, por tanto, se sigue formando una réplica de la escena 3D. Por otro lado, si las diferencias entre el sistema de captura y el monitor de InI tienen su origen en alteraciones de cualquier otro parámetro, los haces ya no se cruzarán en los puntos de la escena que los habían originado en el proceso de captura, sino en otros diferentes. Por tanto, aunque se sigue formando una imagen 3D, ésta ya no será una réplica de la escena 3D original.

Tal y como se analizó en los estudios realizados por Martínez-Corral y colaboradores [50], que los haces de luz no se crucen en los puntos que los originaron puede producir una distorsión en la imagen 3D generada. Por ejemplo, como se puede ver en la Figura 4.4, que la separación entre los píxeles y las lentes del monitor de InI sea diferente a la separación existente entre el sensor y las lentes del sistema de captura, produce una distorsión axial en la imagen 3D.

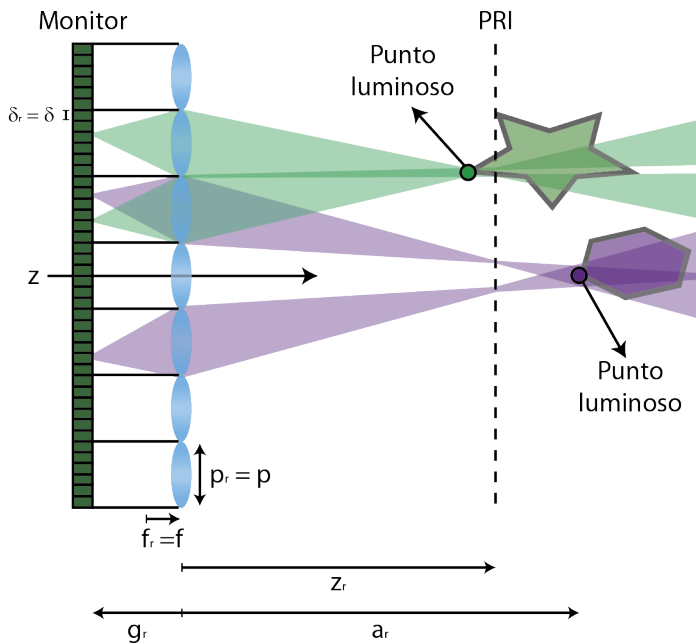


Figura 4.4.: Ejemplo de la imagen 3D distorsionada que produce un monitor de InI en configuración asimétrica con respecto al sistema de captura de la Figura 4.2 (a). En este caso, se ha modificado la distancia g_r para que no sea igual a la distancia g empleada en la captura.

Afortunadamente, la distorsión puede ser reducida o incluso eliminada. Específicamente, en [50] se concluye que la distorsión de la imagen 3D podía ser eliminada si, en el monitor de InI, se escalan por igual los siguientes parámetros: la separación entre los píxeles y las lentes de la MR, g_r , el tamaño de las lentes de la MR, p_r , y el período de captura, T_r (separación en el monitor entre las distintas imágenes de un punto de la escena 3D). En este caso, el monitor de InI produce una imagen sin distorsión y

escalada en el mismo factor que los tres parámetros antes mencionados. No obstante, con el objetivo de aprovechar de manera efectiva los píxeles del monitor cuando se utilizan matrices de cámaras como sistema de captura, Martínez-Cuenca también propone, en un estudio posterior [56], una configuración de monitor de InI alternativa que tampoco produce distorsión en la imagen 3D. De su propuesta podemos deducir que el requisito mínimo para que un monitor de InI produzca imágenes 3D sin distorsión es que el ángulo que subtienden las imágenes que se muestran en él con respecto a los centros de las lentes de la MR, γ_r , sea el mismo que las imágenes capturadas subtendían con respecto a los centros de las lentes en el proceso de captura, γ , lo que se puede expresar de la siguiente forma:

$$\frac{\gamma}{\gamma_r} = \frac{L/g}{L_r/g_r}, \quad (4.2)$$

donde L y L_r son, respectivamente, el tamaño de las imágenes en el sistema de captura y el tamaño de las imágenes en el monitor de InI, ver Figuras 4.5 y 4.6. En consecuencia, cuando un monitor de InI trabaja en configuración asimétrica, la manera más sencilla y efectiva de producir una imagen 3D sin distorsión es ajustando las imágenes capturadas al tamaño de las lentes de la MR y seleccionar la distancia g_r para que cumpla la ecuación 4.2. Hay que tener en cuenta que, en este caso, la imagen 3D reconstruida se encontrará escalada un factor $M = p_r/p$. En la Figura 4.5 se muestra la captura del flujo de luz de una escena usando un dispositivo de InI y su posterior reconstrucción, sin distorsión, con un monitor de InI operando en configuración asimétrica. Del mismo modo, en la Figura 4.6 se muestra la captura del flujo de luz de una escena con un dispositivo de fotografía plenóptica y su posterior reconstrucción, sin distorsión, mediante un monitor de InI operando en configuración asimétrica.

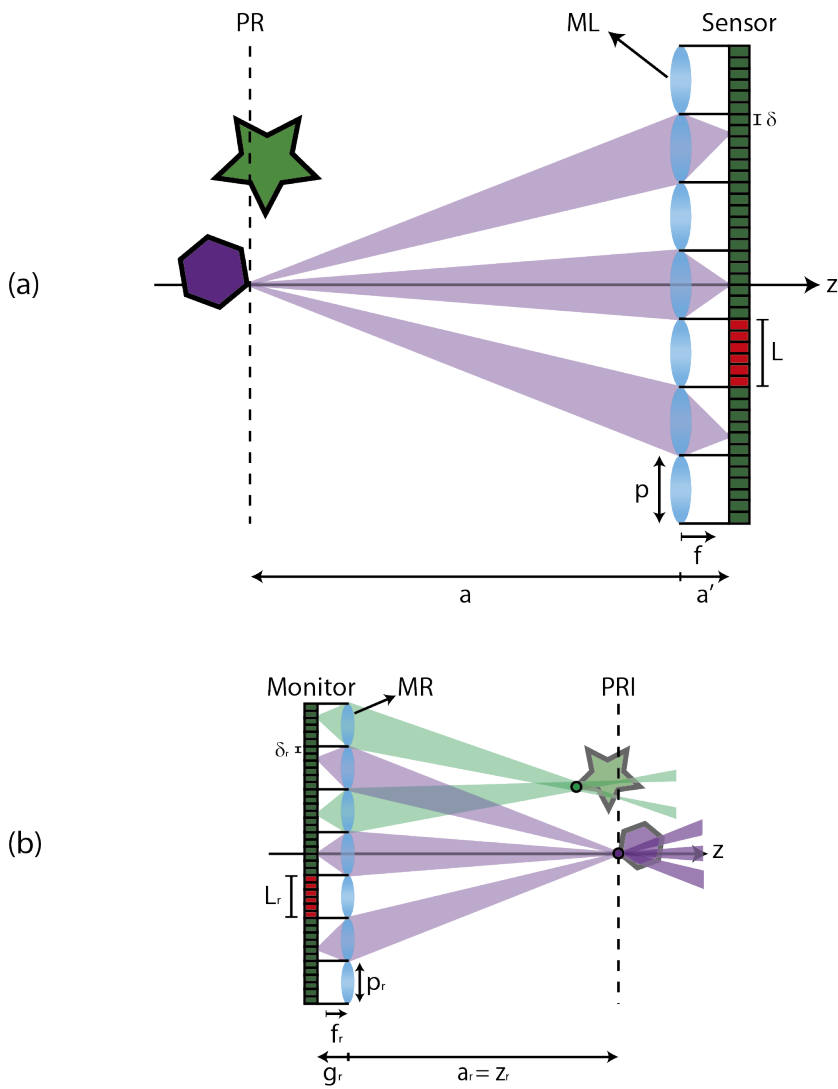


Figura 4.5.: (a) Configuración de captura de InI que emplea una ML y un sensor. (b) Reconstrucción del flujo de luz capturado en (a) mediante un monitor de InI en configuración asimétrica que se encuentra escalado en un factor 0,66 con respecto al sistema de captura.

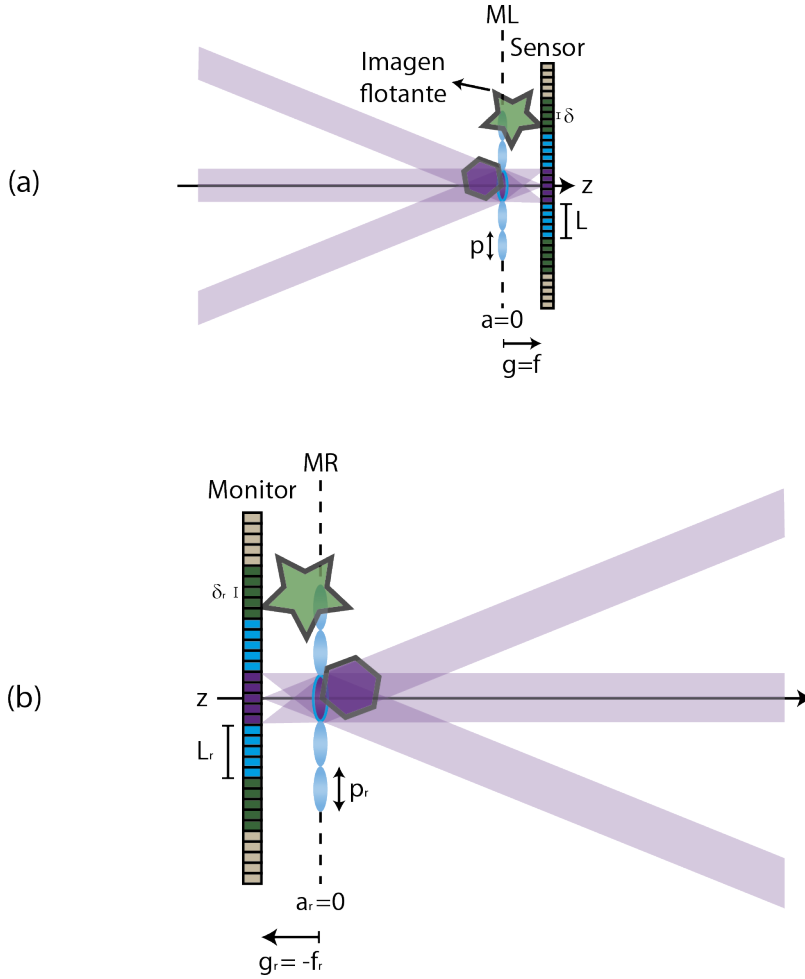


Figura 4.6.: (a) Configuración de captura de fotografía plenóptica. (b) Reconstrucción del flujo de luz capturado en (a) mediante un monitor de InI en configuración asimétrica que se encuentra escalado en un factor 0,66 con respecto al sistema de captura.

4.1.2 Características de la imagen 3D generada

Hasta ahora, hemos visto como la semejanza entre el sistema de captura de la función plenóptica y el monitor de InI afecta a la reconstrucción del flujo de luz y, por tanto, a la imagen 3D generada. En este apartado, comentaremos con más detalle las características de la imagen 3D.

La primera característica que merece ser destacada es que el tamaño y la posición de la imagen 3D que se forma dependen de la configuración en la que trabaja el monitor de InI, tal y como vimos en la sección anterior. Asimismo, si observamos las imágenes 3D generadas en las Figuras 4.2-4.6, vemos que, independientemente de la configuración empleada en el monitor de InI, son pseudoscópicas, es decir, se encuentran invertidas en profundidad con respecto a la escena 3D capturada.

Otra característica importante de la imagen 3D que produce un monitor de InI es su resolución espacial, entendida como el tamaño de los puntos luminosos que la constituyen. Este tamaño depende de la anchura de los haces de luz que, al cruzarse, forman los puntos luminosos. Como la anchura de los haces de luz cambia con la distancia, la resolución espacial de la imagen 3D no es igual en todos sus planos, si no que depende de la distancia, a_r , de cada plano de la imagen 3D a la MR y se puede expresar de la siguiente manera:

$$\eta = \left| \frac{\delta_r a_r}{g_r} \right| + p_r \left| \frac{a_r - z_r}{z_r} \right|, \quad (4.3)$$

donde p_r es el tamaño de las lentes de la MR, δ_r es el tamaño de los píxeles del monitor, g_r es la separación entre estos y las lentes de la MR y z_r es la distancia entre la MR y el PRI. En la Figura 4.7 se muestra cómo varía la resolución que tendría la imagen 3D en función de la distancia de ésta a la MR. A partir de la ecuación 4.3 y la Figura 4.7, podemos ver que, tanto el tamaño de los píxeles del monitor, como las características de las lentes

afectan considerablemente a la resolución de la imagen 3D. Asimismo, a pesar de que la mejor resolución se obtiene en el PRI, ésta empeora rápidamente conforme nos alejamos de ese plano, por lo que la región dónde se obtiene una imagen 3D con buena resolución es limitada.

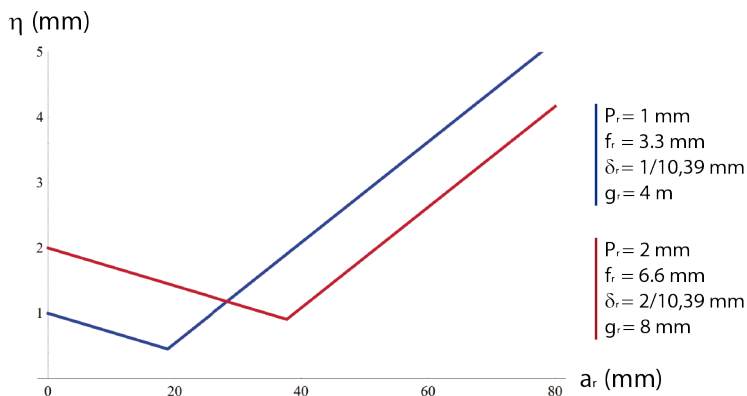


Figura 4.7.: Ejemplo de cómo varía la resolución de la imagen 3D en función de la distancia a la MR. La línea azul representa la resolución de uno de los monitores de InI implementados durante el desarrollo de esta Tesis. La línea roja representa la resolución en el caso de utilizar el mismo monitor pero con una MR escalada en un factor 2.

Otro aspecto que merece ser comentado es el número de vistas diferentes, pertenecientes a la escena original, que contiene la imagen 3D. Por un lado, si la imagen 3D se reconstruye mostrando EIs en el monitor, el número de vistas será igual al número de EIs. Sin embargo, si se muestran microimágenes, el número total de vistas contenidas en la imagen 3D será igual al número de píxeles en cada microimagen. Asimismo, la resolución angular con la que se muestran las vistas depende, en el caso de las EIs, a la separación angular de las lentes respecto al punto central del PR, p_r/z_r , mientras que equivale a δ_r/g_r si se muestran microimágenes.

4.2 Visualización de la imagen 3D

El receptor final de la imagen 3D generada por un monitor de InI es el ojo del observador. Por tanto, es necesario analizar el proceso de

visualización de la imagen 3D por parte del observador para entender cómo se ve afectada la calidad de la imagen cuando ésta es percibida.

Como hemos visto, un monitor de InI forma una imagen 3D por la intersección, en el espacio enfrente de él, de haces de luz. Para visualizar la imagen 3D, un observador situado en frente del monitor de InI ajustará la acomodación y convergencia de su sistema visual sobre los puntos luminosos que se forman en la intersección de estos haces de luz [24], [57], [58]. La existencia de estos puntos luminosos o fuentes de luz es, precisamente, una de las grandes ventajas de la técnica de InI con respecto a otros dispositivos formadores de imágenes 3D, ya que permite a un observador enfocar y converger su sistema visual a los objetos que emiten el flujo de luz que éste o ésta percibe, evitando el conflicto convergencia-acomodación [7]-[10]. Sin embargo, como los ojos no pueden acomodar simultáneamente a todos los planos dónde se encuentran los puntos luminosos que constituyen la imagen 3D, el observador seleccionará un plano, donde esté el objeto o región de interés, y fijará su acomodación sobre él. Los puntos de la imagen 3D situados en este plano, al que denominaremos *plano de visualización* (PV), son los únicos que se encontrarán focalizados en la retina [24], mientras que el resto de la imagen 3D será percibida con cierto desenfoque. No obstante, dada la capacidad de formar imágenes de las lentes de la MR, es posible que se produzca una degradación en la imagen que el observador percibe debido al efecto conocido como *trenzado de facetas*¹ [59], [60].

Como, tal y como vimos en la sección 4.1.1, los haces emitidos por un monitor de InI convergen en el PRI, las regiones de la imagen 3D que se formen fuera del PV del observador, pero cercanas al PRI, serán percibidas como una superposición de pequeñas imágenes parciales y desplazadas [59], [60], ver Figura 4.8. Para evitar esta discontinuidad en la percepción de la imagen 3D, Navarro y colaboradores [60] propusieron situar los píxeles del monitor en el plano focal de las lentes de la MR, $g_r = f_r$. De

¹Se conoce como *faceta* a la porción de la imagen 3D que el observador visualiza a través de cada lente.

este modo, las lentes ya no forman imágenes y, por tanto, el observador no percibirá imágenes parciales de los objetos o regiones sobre los que no focalice, sino una versión desenfocada de los mismos. La resolución espacial de la imagen reconstruida al aplicar esta modificación empeora de manera constante con la distancia, y se expresa como:

$$\eta = \left| \frac{\delta_r a_r}{g_r} \right| + p_r. \quad (4.4)$$

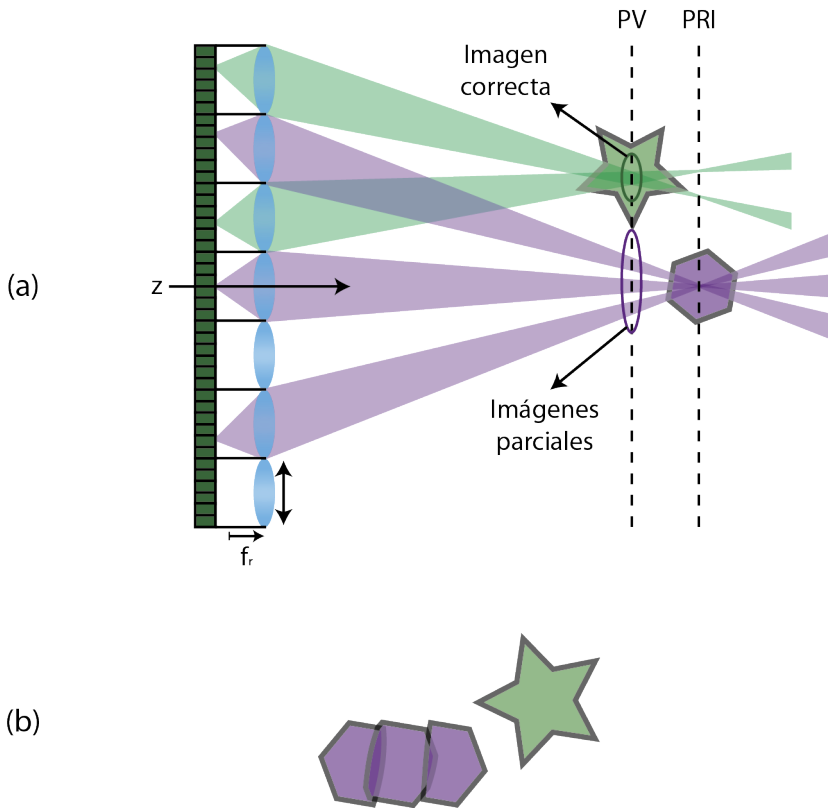


Figura 4.8.: (a) Generación de la imagen 3D por un sistema de InI. (b) Imagen percibida por el observador.

A parte de eliminar el efecto negativo del trenzado de facetas, esta modificación permite producir una imagen 3D con mejor resolución para planos de la imagen alejados del PRI. En la Figura 4.9 se compara la

resolución que se obtiene en el caso convencional, ecuación 4.3, con la resolución obtenida con la modificación propuesta por Navarro y colaboradores [60], ecuación 4.4. Se puede apreciar que la resolución en las zonas cercanas al PRI empeora si se realiza esta modificación. Sin embargo, la resolución es mejor para el resto de los planos y se degrada más lentamente que en el caso convencional. En consecuencia, al situar los píxeles del monitor en el plano focal de las lentes de la MR, la región dónde se produce una imagen 3D con resolución aceptable es mayor que en el caso convencional.

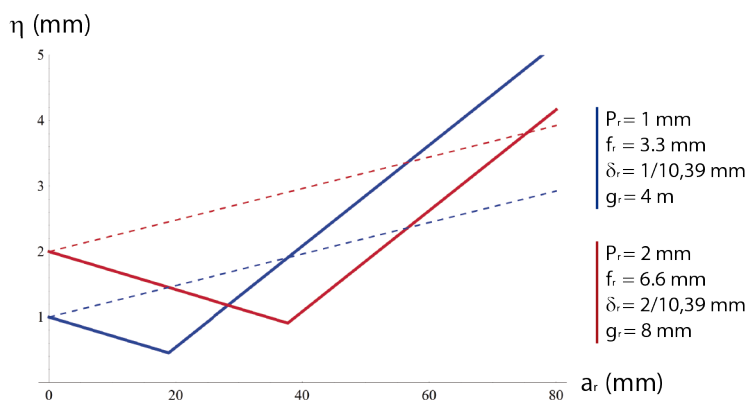


Figura 4.9.: Comparación entre la resolución de la imagen 3D en función de a_r de los ejemplos mostrados en la Figura 4.7 (línea continua) y esos mismos ejemplos pero con un valor de $g_r = f_r$ (línea discontinua).

Desafortunadamente, realizar esta modificación introduce un par de problemas. El primero es debido a la alteración de uno de los parámetros del monitor de InI. Forzar que el valor de g_r sea igual a f_r puede provocar una distorsión axial en la imagen 3D reconstruida. Sin embargo, esta distorsión sería imperceptible para un observador en el caso de que el nuevo valor de g_r no sea muy diferente al valor original [24]. El segundo problema es consecuencia de situar los píxeles del monitor en el plano focal de las lentes de la MR. En este caso, las lentes forman una imagen de la estructura de los píxeles de la pantalla en el infinito. En consecuencia, un observador situado frente al monitor percibirá la imagen de los píxeles superpuesta a la imagen 3D formada por el monitor de InI. En la Figura

4.10 se puede apreciar como este fenómeno, al que denominaremos efecto subpixel, empeora considerablemente la calidad de la imagen percibida. Es interesante mencionar que existen distintas estructuras de píxeles, ver Figura 4.11. Los píxeles de las pantallas que se utilizan normalmente están divididos en subpíxeles de los tres colores primarios: rojo, verde y azul. La combinación de la luz procedente de los tres subpíxeles es lo que le permite al pixel emitir luz del color deseado. Hay que destacar que, como la imagen de los subpíxeles es un patrón periódico, éste puede interactuar con la estructura periódica de la MR, creando un patrón de Moiré [61]-[63].



Figura 4.10.: Monitor de InI en el que $g_r = f_r$.

A pesar de los problemas que ocasiona, en esta Tesis consideramos que, siempre que sea posible, es necesario aplicar la modificación propuesta por Navarro y colaboradores [60] para eliminar el efecto del trenzado de facetas. No obstante, se puede realizar un ligero desplazamiento en la posición de los píxeles del monitor para que estos no se encuentren justamente en el plano focal de las lentes [39], [64]. De este modo, es posible reducir la visibilidad de la estructura de los píxeles sin acentuar excesivamente el efecto del trenzado de facetas.

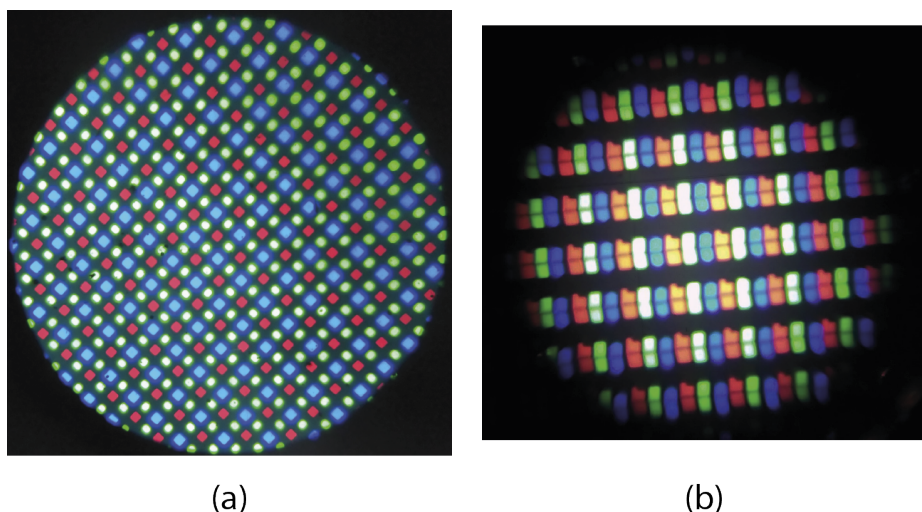


Figura 4.11.: Imagen de los píxeles de una pantalla vistos con un microscopio.
 (a) Pantalla con estructura subpixel en forma de panel de abeja;
 (b) Pantalla con estructura subpixel en forma rectangular.

4.2.1 Área de visualización de un monitor de InI

Otro importante factor a tener en cuenta en el proceso de visualización es la posición del observador con respecto a la pantalla. Un monitor de InI tiene una zona en frente donde un observador podrá percibir una imagen 3D de forma correcta. Denominamos a esta zona como el *área de visualización* (AV) del monitor de InI. Si los dos ojos del observador se encuentran en el interior de esta AV, percibirá la imagen 3D con paralaje completo. Si el observador se mueve fuera de esta zona, percibirá una imagen 3D distorsionada. Esta distorsión ocurre debido a que, en este caso, algunos de los haces de luz recibidos por el observador, y que han sido emitidos por el monitor de InI, provienen de lentes incorrectas. Una representación esquemática de este efecto, conocido como crosstalk, se muestra en la Figura 4.12. Como se aprecia en el esquema, algunos de los haces de luz que recibe el observador a través de algunas de las lentes no han sido emitidos por sus correspondientes imágenes, por tanto, percibe información incorrecta. En la Figura 4.13 se muestra una

fotografía realizada en frente de un monitor de InI con la cámara en una posición situada fuera del AV. En esta imagen se puede observar como el crosstalk causa una distorsión en algunas de las zonas de la imagen percibida por el observador, como en el tejado y la puerta de la casa y, también, en la parte derecha del bloque situado en la parte frontal de la escena.

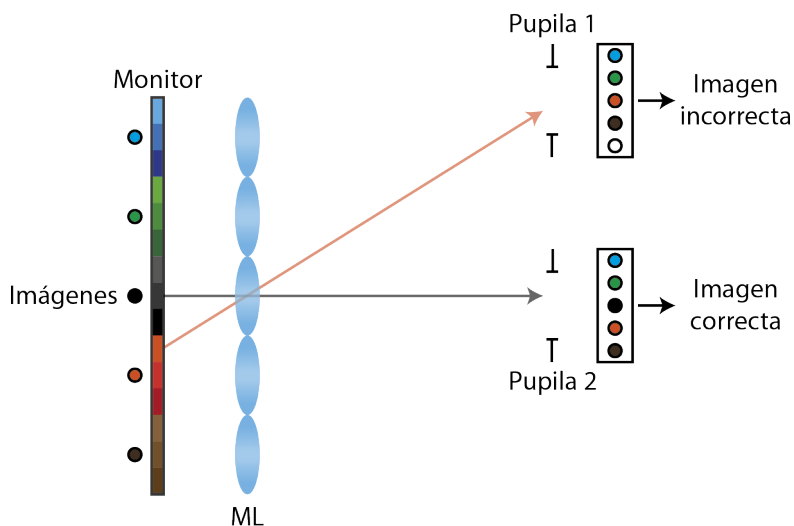


Figura 4.12.: Representación esquemática del efecto crosstalk. Sólo un ojo del observador recibe la información correcta de la lente central (color negro).

Para obtener el AV de un monitor de InI, debemos encontrar las posiciones en las cuales un observador recibe la información correcta a través de todas las lentes. Esto se consigue al proyectar cada una de las imágenes a través del centro de su correspondiente lente en el espacio que hay frente a ellas. El área común de todas las proyecciones define el AV del monitor de InI, ver Figura 4.14. A una distancia de la MR determinada, la proyección de una imagen, I_{proy} , y el AV, pueden ser calculadas como:

$$I_{Proy}(z) = \frac{p_r z}{g_r}, \quad (4.5)$$

$$AV(z) = \frac{p_r z}{g_r} - (l - 1)p_r - \mu, \quad (4.6)$$

donde l es el número de imágenes que se muestran en el monitor de InI, μ es la separación inter-pupilar y z es la distancia de proyección medida desde la MR. Si nos fijamos en las ecuaciones 4.5 y 4.6, y en la Figura 4.14, podemos ver que el número de imágenes que se muestran en el monitor de InI limita el AV. Por tanto, cuanto mayor sea su número, más pequeña será el AV.



Figura 4.13.: Ejemplo de la distorsión que produce el crosstalk en la imagen percibida.

Es conveniente mencionar también que, como comentamos en la sección 4.1.1, a no ser que se introduzcan barreras para separar las distintas imágenes que se muestran en el monitor de InI, se generan flujos de luz secundarios cuando los haces de luz de las imágenes atraviesan lentes no correspondientes. Estos flujos de luz secundarios formarán imágenes 3D distorsionadas en distintas posiciones. No obstante, la distorsión

producida es pequeña y suele ser imperceptible para el observador. En consecuencia, el AV de un monitor de InI puede no ser única. La proyección de las imágenes a través de lentes adyacentes, en lugar de a través de las correspondientes, dará lugar a áreas de visualización secundarias, ver Figura 4.15. En estas áreas oblicuas con respecto al monitor de InI, el observador percibiría correctamente la imagen 3D.

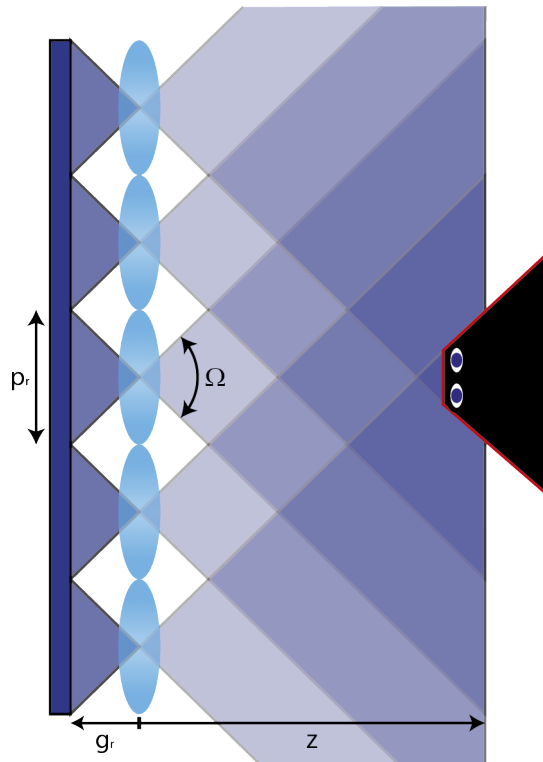


Figura 4.14.: Representación esquemática del AV de un monitor de InI.

Un parámetro complementario al AV es el ángulo de visualización del monitor de InI. Este parámetro se define sólo para un ojo u observador monocular. Este ángulo se encuentra formado por la normal a la lente central del monitor con respecto a la línea que conecta el centro de la pupila de entrada del ojo de un observador, situado en un extremo de un AV, con el centro óptico de la lente central. Como el ojo del observador

debe encontrarse en el interior del AV, el ángulo de visualización suele ser pequeño.

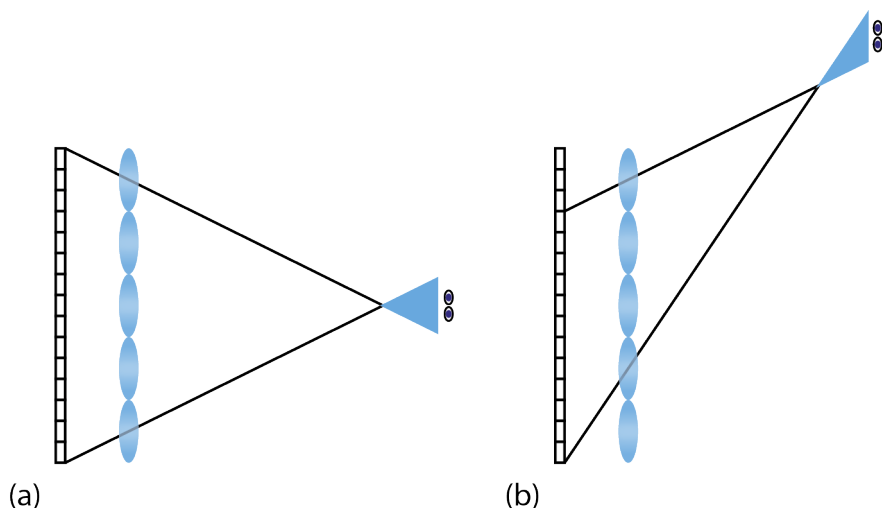


Figura 4.15.: (a) AV principal de un monitor de InI; (b) Una de las áreas de visualización secundarias del mismo.

4.2.2 Resolución de la imagen percibida

Para finalizar este análisis del proceso de visualización de la imagen 3D, veremos como la resolución percibida por el observador suele ser diferente a la de la imagen 3D generada.

En general, no toda la luz que procede de una lente de la MR llega a la retina. Si nos fijamos en la Figura 4.9, podemos ver que, en los casos de ejemplo, la anchura del haz emitido por un pixel del monitor de InI, es decir, la imagen de un pixel en planos alejados de la MR, será de un tamaño mayor que el diámetro de la pupila de un observador, ver Figura 4.16. Este hecho tiene dos consecuencias. La primera es que cada uno de los ojos de un observador sólo recibirá luz procedente de un pixel de cada imagen, es decir, un observador sólo “verá” un pixel tras cada lente de la MR. La segunda es que la resolución de la imagen 3D que el observador percibe puede ser diferente a la de la imagen 3D que se

genera (ecuaciones 4.3 y 4.4). Por tanto, la resolución percibida por un observador se expresa:

$$\eta_v = \text{Min} \left\{ \left| \frac{\delta_r a_r}{g_r} \right| + p_r \Delta, p_r + \frac{\epsilon a_r}{z} \right\} \quad (4.7)$$

donde Min es el valor mínimo entre las dos expresiones en el interior del corchete, z es la distancia del ojo a la MR, ϵ representa la diferencia entre el tamaño de la pupila del ojo y el tamaño de las lentes de la MR y Δ es un término que depende de la distancia entre los píxeles del monitor y la MR. Si los píxeles del monitor se encuentran en la focal de las lentes de la MR, Δ tendrá un valor de 1, en cualquier otro caso tomará la expresión $\Delta = (z - z_r)/z_r$. Analizando esta ecuación podemos deducir que, en general, la resolución de la imagen 3D percibida por un observador vendrá delimitada por el tamaño de las lentes de la MR.

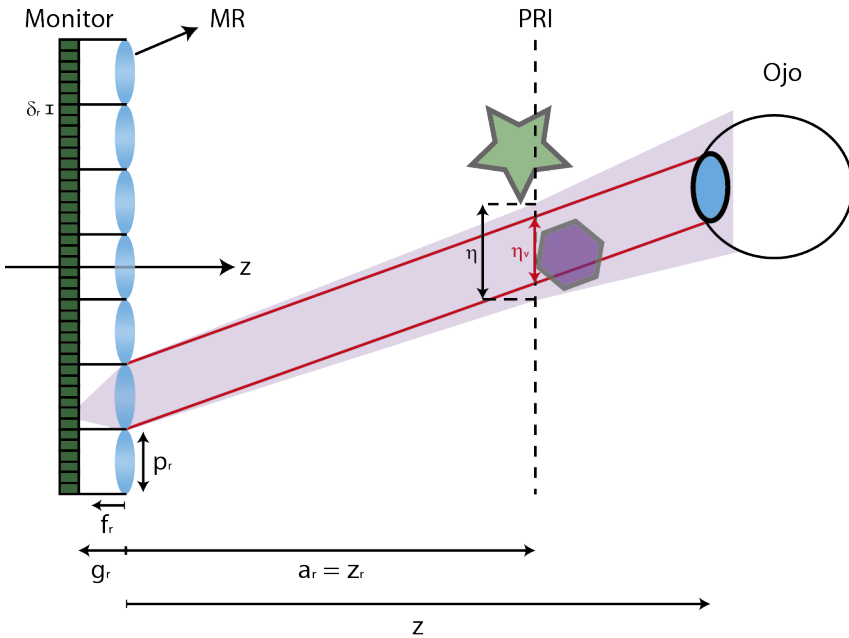


Figura 4.16.: Representación esquemática del efecto que produce la pupila en la percepción de la imagen 3D por parte del observador.

4.2.3 Propuesta de monitor de InI

A continuación, teniendo en cuenta todo lo tratado en este capítulo, se plantean las características generales que debería presentar un monitor de InI para que las imágenes 3D que genere sean percibidas con cierta calidad por parte del observador.

En concreto, queremos que el monitor de InI tenga la capacidad de generar imágenes 3D con buena resolución y gran paralaje de escenas extensas, compuestas por varios objetos. Sin embargo, la resolución de toda la imagen 3D debe ser aceptable, no queremos que se reconstruya con buena resolución sólo uno de los objetos o planos que la componen y que el resto de la imagen se muestre excesivamente desenfocada.

De acuerdo con la ecuación 4.7 y las Figuras 4.9 y 4.16, la resolución de la imagen 3D percibida por el observador depende principalmente del tamaño de las lentes de la MR. Por tanto, un monitor de InI debe utilizar lentes lo más pequeñas posible (teniendo cuidado de no aumentar en exceso los efectos difractivos). Además, para reducir la pérdida de resolución ocasionada por el efecto del trenzado de facetas, los píxeles deberán situarse cerca del plano focal de las lentes, lo que, como vimos en la sección 4.2, permitirá generar una imagen 3D extensa en la que no decaiga rápidamente la resolución. En estas condiciones, la mejor resolución se obtiene en las cercanías de la MR, por tanto, nos interesa que la imagen 3D se genere cerca de ella. Hay que destacar que para aprovechar al máximo el tamaño de la pantalla del monitor, la MR empleada deberá tener un número suficiente de lentes para poder abarcarla por completo. Asimismo, como la separación entre los píxeles del monitor y las lentes influye en el tamaño del AV, interesa que la focal de las lentes sea pequeña pero sin que se produzcan distorsiones.

Teniendo en cuenta todo lo comentado en el párrafo anterior, resulta evidente que al mostrar microimágenes obtenidas con una cámara ple-

nóptica en el monitor de InI, se obtendrá una imagen 3D con la mejor resolución posible. En este caso, la imagen 3D se genera sobre la MR² y el número de vistas diferentes de la imagen dependerá del número de píxeles por microimagen. Como el tamaño de las microimágenes depende del tamaño de las lentes, se requiere cierto compromiso entre el valor de ambos parámetros para obtener una imagen 3D con un número de vistas adecuado sin que por ello empeore excesivamente su resolución. Por ello, conviene utilizar un monitor de alta resolución, con los píxeles lo más pequeños posible.

²Según nuestra experiencia, que haya partes de la imagen 3D por delante y por detrás de la MR proporciona una mejor experiencia de visualización al observador que si ésta se reconstruye toda por delante o toda por detrás

Método para la generación de microimágenes adaptadas a un monitor de InI

En el capítulo anterior planteamos las características de debería tener un monitor de InI para que generara imágenes 3D de calidad. Sin embargo, obtener las microimágenes que se van a mostrar en él tiene algunas dificultades importantes. Por un lado, las microimágenes que se capturen necesitarán ser adaptadas para poder ser mostradas en el monitor de InI. Por otro, los sistemas convencionales de obtención de microimágenes presentan varias limitaciones que dificultan, tanto la obtención de las microimágenes adecuadas que se van a mostrar en el monitor de InI, como la generación de imágenes 3D de escenas grandes, extensas y que contengan un gran paralaje. Asimismo, la inversión en profundidad que sufren las imágenes 3D producidas por un monitor de InI empeora la experiencia de visualización por parte de un observador.

En este capítulo, profundizaremos en estos problemas y se propondrá un método que puede resultar de utilidad para minimizarlos o, incluso, evitarlos [65]. El método se basa en la relación de equivalencia entre los sistemas de captura de InI y fotografía plenóptica que vimos en la sección 3.4. Nuestra propuesta permite obtener microimágenes listas para ser mostradas en un monitor de InI y que producen imágenes 3D ortoscópicas de escenas grandes, extensas y con gran paralaje. Además, también ofrece

la posibilidad de seleccionar la resolución, campo visual y posición de la imagen generada.

5.1 Adaptación de las microimágenes a un monitor de InI

Como vimos en la sección 4.1.1, es difícil obtener las microimágenes que se van a mostrar en un monitor de InI con un sistema de captura que sea una réplica de éste. Por tanto, es habitual que el monitor de InI use un sensor y una MR de características diferentes a sus homólogos empleados en el proceso de captura, es decir, que trabaje en modo asimétrico. En este caso, la obtención de las microimágenes adecuadas no es una tarea sencilla, ya que el número de microimágenes capturadas, $N_H \times N_V$, y/o su número de píxeles, $n_H \times n_V$, será diferente al número de microimágenes, $M_H \times M_V$, y/o píxeles por microimagen, $m_H \times m_V$, que el monitor de InI puede mostrar, ver Figura 5.1. Bajo estas circunstancias, será necesario modificar las microimágenes capturadas para adaptarlas al monitor de InI¹.

Podemos dividir el proceso de adaptación en dos partes. La primera parte consiste en asegurarse de que la imagen 3D que se genera sea correcta. Para ello, tal y como vimos en la sección 4.1.1, es necesario ajustar el tamaño de las microimágenes obtenidas al tamaño de las lentes de la MR, $L = L_r$. Esto puede realizarse aplicando algoritmos de interpolación directamente a las microimágenes para aumentar o reducir su número de píxeles. Por otro lado, la segunda parte consiste en ajustar el número de microimágenes para que coincida con el número que deseamos mostrar. Lo ideal sería aprovechar al máximo la pantalla del monitor de InI y, por tanto, ajustar el número de microimágenes al número de lentes de la MR.

¹La única excepción a esta situación ocurre cuando el sistema de captura es una réplica escalada del monitor de InI. No obstante, esta configuración es igual de difícil de implementar que la configuración simétrica.

Este ajuste también puede hacerse con algoritmos de interpolación pero, en este caso, aplicarlos directamente sobre las microimágenes no serviría para cambiar su número. En su lugar, la interpolación deberá aplicarse sobre las subimágenes.

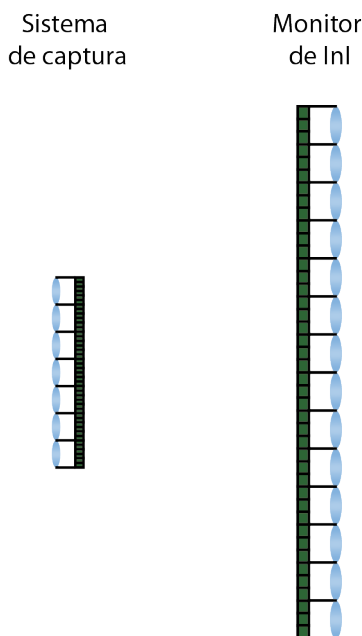


Figura 5.1.: Sistema de captura y monitor de InI con distinto número de lentes y píxeles tras cada lente.

Como vimos en la sección 3.3, las subimágenes se obtienen extrayendo el mismo píxel de cada microimagen, siendo el número de píxeles de las subimágenes dependiente del número de microimágenes de la imagen plenóptica. Por tanto, para alterar el número de microimágenes capturadas debemos, en primer lugar, obtener las subimágenes. Después, aplicaremos algoritmos de interpolación para cambiar el número de píxeles por subimagen de tal forma que éste sea igual al número de microimágenes deseado. Y, por último, realizaremos la transformación inversa para convertir las subimágenes de nuevo en microimágenes. Hay que destacar que ambos pasos del proceso de ajuste también pueden realizarse interpolando directamente sobre la función plenóptica codificada en las microimágenes. Como vimos en la sección 3.3.1, una columna de puntos en el diagrama

espacio-angular de la función plenóptica capturada representa una microimagen, mientras que cada fila representa una vista o subimagen diferente de la escena 3D. Por tanto, podemos alterar el número de microimágenes y/o el número de vistas al añadir, o eliminar, columnas y/o filas de puntos en el espacio de rayos de la función plenóptica mediante algoritmos de interpolación, ver Figura 5.2. No obstante, hay que tener en cuenta que interpolar directamente en la función plenóptica es un proceso más costoso computacionalmente, ya que es una función 4D.

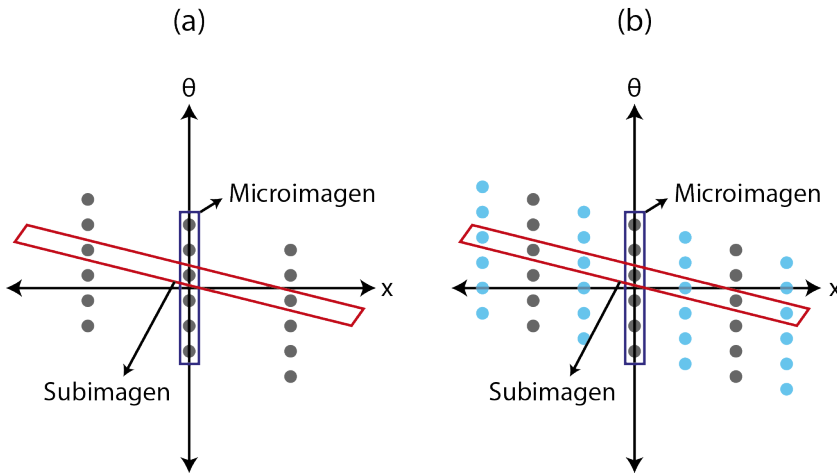


Figura 5.2.: Ejemplo de interpolación de la función plenóptica en el que se aumenta el número de microimágenes añadiendo columnas de puntos. (a) Función plenóptica capturada. (b) Función plenóptica tras la interpolación.

Es importante remarcar que una interpolación no añade información nueva. Por tanto, si en el proceso de adaptación se han añadido microimágenes y/o aumentado el número de píxeles por microimagen, la resolución espacial y/o angular de la imagen 3D no se corresponderá con la que el monitor de InI puede producir (ver sección 4.1.2) sino que será peor. Debido a ello, cuando el número de microimágenes disponibles, $N_H \times N_V$, y/o su número de píxeles, $n_H \times n_V$, sea menor al número de microimágenes, $M_H \times M_V$, y/o píxeles por microimagen, $m_H \times m_V$, que el monitor de InI puede mostrar, no se estará aprovechando correctamente las características del monitor de InI, es decir, las microimágenes obtenidas no

serán adecuadas. En consecuencia, es evidente que para generar la mejor imagen 3D posible, el sistema de captura debe obtener un número mayor o igual de microimágenes y de píxeles por microimagen al que puede ser mostrado en el monitor de InI. De este modo, incluso si es necesario realizar el proceso de ajuste, la resolución espacial y angular de la imagen 3D que se genera se corresponderá con la resolución que el monitor de InI puede producir.

Teniendo en cuenta todo lo visto hasta el momento, queda claro que el proceso de adaptación de las microimágenes obtenidas requiere tiempo, lo que supone un inconveniente a la hora de implementar un monitor de InI que funcione en tiempo real.

Ejemplo de adaptación

A continuación, a modo ilustrativo, se describe el proceso de adaptación de un conjunto de microimágenes a un monitor de InI.

En primer lugar, necesitamos capturar las microimágenes y, para ello, usaremos una cámara plenóptica. A pesar de que ya existen versiones comerciales [66], [67], decidimos preparar en nuestro laboratorio una cámara plenóptica en configuración abierta, ver Figura 5.3 y Figura 5.4. En esta configuración, se usó una lente de focal $f_L = 100$ mm para conjugar un plano de la escena, el PR, con el plano de la ML. La ML estaba compuesta por 96×60 lentes de focal $f = 0,93$ mm y tamaño $p = 0,222$ mm, distribuidas en arreglo cuadrado (APO-Q-P222-F0.93 de AMUS). Una cámara digital (Canon 450D) con un objetivo macro 1:1 se utilizó como vehículo para que las imágenes se formaran sobre el sensor [56]. La imagen obtenida con esta cámara plenóptica se muestra en la Figura 5.5 (a). El número de microimágenes capturadas fue de 96×60 , con 48×48 píxeles cada una.

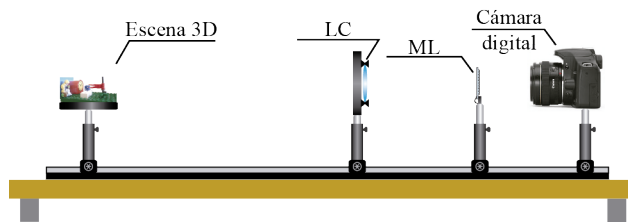


Figura 5.3.: Esquema de una cámara plenótica en configuración abierta.

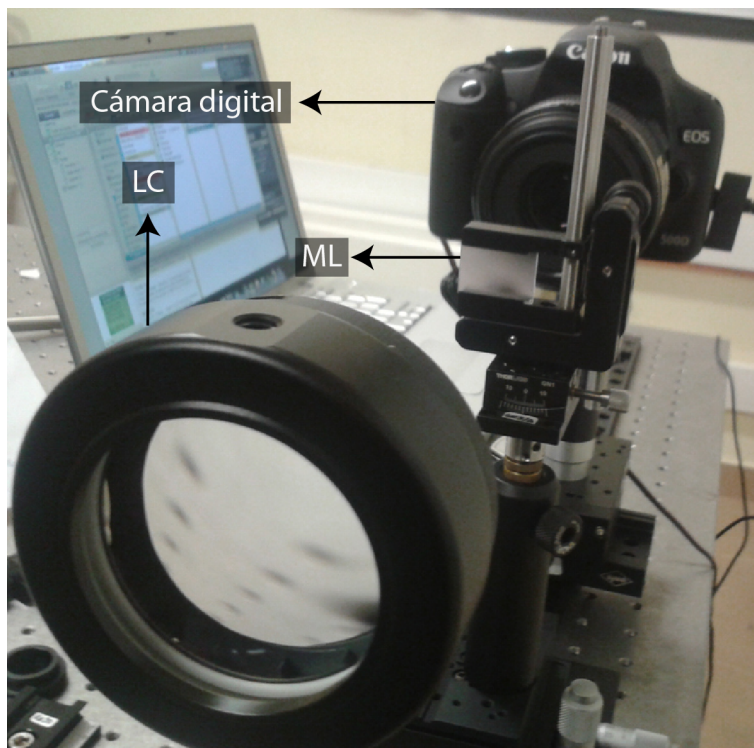


Figura 5.4.: Fotografía de la cámara plenótica empleada en la captura.

Para el monitor de InI usamos un iPad con pantalla retina (resolución de 10,39 píxeles/mm), y una MR compuesta por lentes de focal $f_r = 3,3$ mm y tamaño $p_r = 1$ mm (Modelo 630 de Fresnel Technology). Este monitor de InI, que se observa en la Figura 5.5 (b), puede mostrar, como máximo, 147×147 microimágenes de $10,39 \times 10,39$ píxeles cada una.

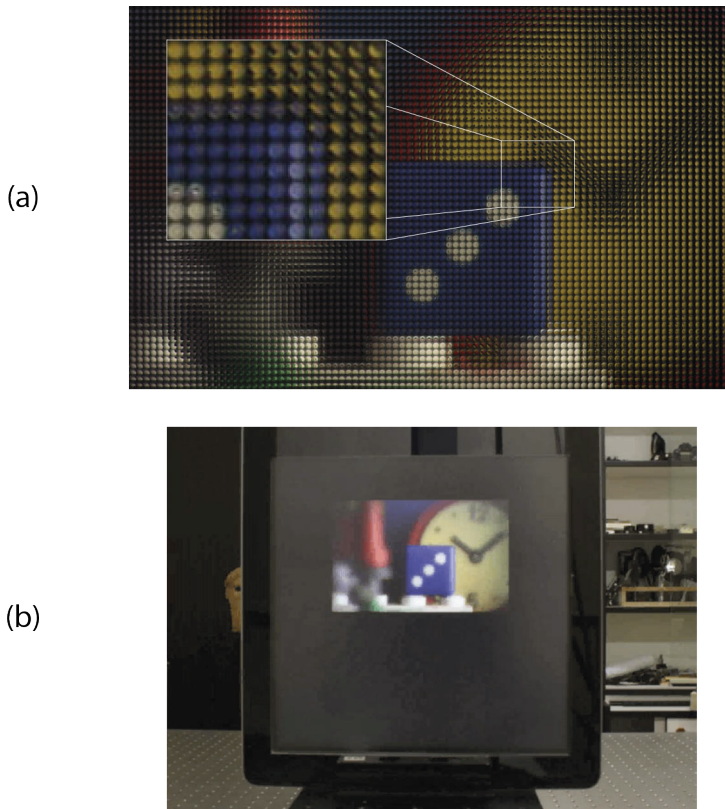


Figura 5.5.: (a) Imagen plenótica de una escena 3D capturada con una cámara plenótica. La imagen está compuesta por 96×60 microimágenes con 48×48 píxeles cada una. (b) Monitor de InI que, como máximo, puede mostrar 147×147 microimágenes de $10,39 \times 10,39$ píxeles cada una.

Para adaptar la imagen plenótica capturada, cambiamos el número de píxeles de la imagen plenótica de 4608×2880 a 997×623 píxeles. Después, para reducir lo máximo posible el efecto del trenzado de facetas sin producir un deterioro de la imagen percibida debido al efecto subpixel

(ver sección 4.2), alineamos la MR en frente del iPad para que los píxeles se encontraran situados cerca del plano focal de las lentes. Hay que destacar que el iPad tiene una placa transparente que cubre los píxeles, lo que limita la separación que puede haber entre los píxeles y la MR (ver Anexo A). Nótese que, cómo el número de píxeles de las microimágenes que se pueden mostrar en el monitor de InI no es entero, el tamaño de éstas no se ajustará perfectamente al tamaño de las lentes. No obstante, esta variación es tan pequeña que la distorsión que ocasiona es imperceptible para un observador.

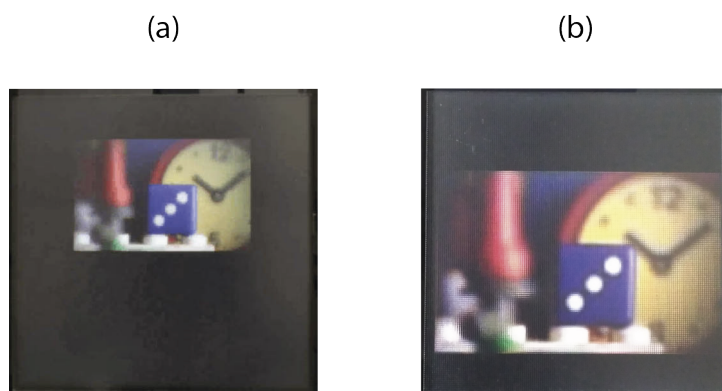


Figura 5.6.: Monitor de InI mostrando: (a) la imagen plenóptica capturada; (b) imagen plenóptica interpolada.

En la Figura 5.5 (b) se muestra el monitor de InI con la imagen que hemos adaptado. Asimismo, hemos creado una película en la que simulamos el proceso de visualización de la imagen 3D por parte de un observador, ver **VÍDEO 5.1**. Con nuestra cámara plenóptica no ha sido un problema obtener el número suficiente de píxeles por microimagen. No obstante, no hemos podido obtener el número adecuado de microimágenes, lo que ha impedido que aprovechemos correctamente la capacidad del monitor de InI. Y esto ocurre aunque aumentemos el número de microimágenes mediante algoritmos de interpolación, ver Figura 5.6, ya que se genera una imagen 3D con una resolución espacial inferior a la que podría obtenerse si se hubiera capturado el número adecuado de microimágenes.

5.2 Sistemas de captura convencionales

La forma convencional de obtener microimágenes es utilizando los métodos comentados en el capítulo 3, una cámara plenóptica o un sistema de InI, que pueden obtener microimágenes directamente en el sensor del sistema de captura.

Aunque con ninguno de los dos métodos es complicado alcanzar un número igual o superior de píxeles por microimagen al que puede ser mostrado en un monitor de InI con las características planteadas en la sección 4.2.3, no sucede lo mismo con respecto al número de microimágenes capturadas, ya que suele ser difícil obtener un elevado número de ellas. Entre los motivos por los que es difícil capturar suficientes microimágenes, destaca que el tamaño de la ML que se suele emplear en estos sistemas no suele ser muy grande, lo que limita el número de microimágenes que caben en ella. Por ejemplo, como comentamos en la sección 3.5, en una cámara plenóptica no es posible utilizar matrices de lentes mayores que el tamaño del sensor de la cámara, que no suele superar los dos centímetros. Por otro lado, en un sistema de captura de InI, el método habitual de obtención de microimágenes consiste en utilizar una cámara como vehículo para llevar las imágenes al sensor [39], [59], [68], [69]. Este método no es sólo complejo de implementar correctamente si no que, además, también limita el tamaño de la ML que se puede utilizar.

Además del inconveniente que acabamos de comentar sobre los métodos convencionales de obtención de microimágenes, nos encontramos con que el paralaje de la escena 3D que estos pueden obtener es limitado. En concreto, éste se encuentra determinado por el ángulo que, visto desde el centro del PR, subtiende la LC, en el caso de utilizar una cámara plenóptica, y el ángulo que subtiende la ML, en el caso de utilizar un sistema de InI, ver Figura 5.7. Por tanto, solamente en el caso de fotografiar escenas 3D pequeñas y cercanas, podrán estos sistemas capturar microimágenes que proporcionen una imagen 3D que tenga un paralaje satisfactorio.

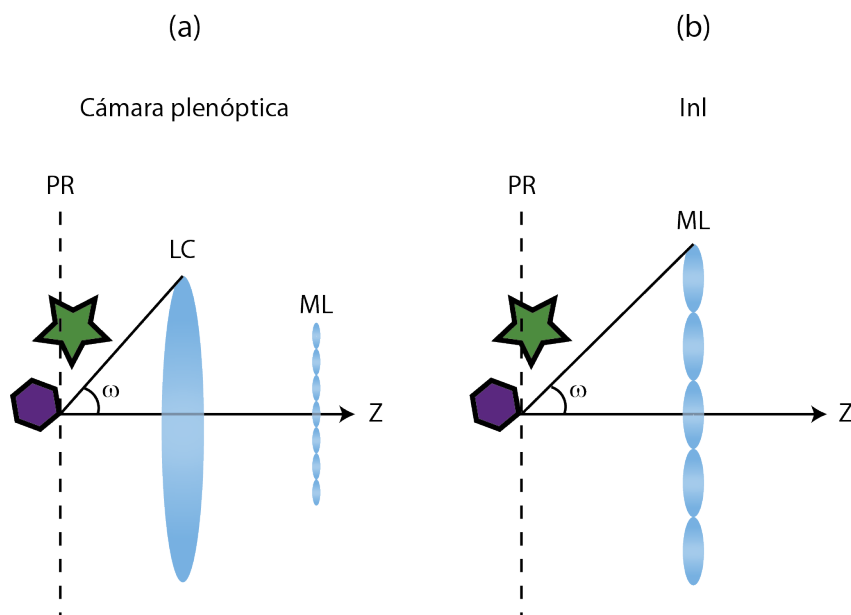


Figura 5.7.: Esquema en el que se muestra el ángulo, ω , que determina el paralaje obtenido en una cámara plenóptica, imagen (a), y el obtenido en un sistema de InI, imagen (b).

5.3 Imágenes pseudoscópicas

Como vimos en la sección 4.1.2, las imágenes 3D que generan los monitores de InI son pseudoscópicas, es decir, se encuentran invertidas en profundidad. Este hecho deteriora la experiencia de visualización de un observador y, por ello, es conveniente corregir este efecto.

El primero en analizar el problema de las imágenes pseudoscópicas fue H. E. Ives [18], que propuso como solución un proceso de captura y reconstrucción de doble etapa. Su método consiste en realizar una captura y reconstrucción de la imagen 3D de manera convencional para, posteriormente, utilizar la imagen pseudoscópica así generada como escena de un segundo proceso de captura y reconstrucción. De este modo, la imagen 3D producida en este segundo proceso es ortoscópica, es decir, no se encuentra invertida en profundidad. Desde H. E. Ives [18], se han

propuesto múltiples opciones para corregir la inversión en profundidad que realizan los monitores de InI² [12], [22], [69]-[76]. La mayoría de estos métodos han sido planteados para ser implementados en un monitor de InI que trabaja en modo simétrico y consisten, en general, en procesar las imágenes capturadas antes de ser mostradas en el monitor. Como este procesamiento adicional de las imágenes no ha sido ideado teniendo en cuenta la configuración asimétrica, estos métodos no suelen abordar el proceso de adaptación de las imágenes capturadas al monitor de InI.

No obstante, destaca entre ellos el algoritmo SPOC (smart pseudoscopic-to-orthoscopic conversion) propuesto por Navarro y colaboradores [75]. Mediante la utilización de un proceso de doble etapa similar al propuesto por Ives [18], no sólo se consigue generar una imagen 3D ortoscópica sino que, además, se pueden adaptar las imágenes capturadas a un monitor de InI que trabaja en modo asimétrico. Su método consiste en realizar una captura con un sistema de InI para, posteriormente, simular de forma computacional un proceso de reconstrucción de la imagen 3D. Esta imagen 3D sintética es después utilizada como escena para una segunda captura de InI, también computacional. Las características del sistema empleado en la captura sintética se escogen para que coincidan con las del monitor de InI real que se empleará para generar la imagen 3D final. De este modo, las microimágenes que se obtienen en esta segunda captura se encuentran ya adaptadas al monitor de InI. Desafortunadamente, durante el proceso de captura virtual del método SPOC se pierde parte de la información plenóptica capturada.

Una de las partes más interesantes de esta propuesta [75] es cómo se realiza la captura de la información plenóptica. Para ello, se utiliza un sistema de InI que obtiene EIs directamente sobre el sensor, lo que permitiría, en teoría, conseguir un gran paralaje de escenas extensas compuestas por múltiples objetos. No obstante, en el trabajo se utiliza un único objeto sobre fondo negro como escena 3D. Esta elección se

²Un resumen de algunas de estas propuestas se puede leer en Navarro [24].

debe, probablemente, a la pérdida de información que se produce en este método, tal y como se ha comentado anteriormente.

5.4 SPOC 2.0

De acuerdo con lo comentado hasta el momento, queda claro que los métodos convencionales de captura de microimágenes y de corrección del problema pseudoscópico no son los más adecuados. Por tanto, a continuación plantearemos un método alternativo para obtener microimágenes que aprovechen al máximo las características de un monitor de InI que trabaje en modo asimétrico y que además, genere imágenes 3D ortoscópicas y con gran paralaje de escenas 3D grandes.

5.4.1 Regreso a la equivalencia entre InI y fotografía plenóptica

Los sistemas de captura que obtienen, mediante el uso de una matriz de cámaras, EIs (o vistas) directamente sobre su sensor, no sufren los problemas que presentan los métodos convencionales de obtención de microimágenes mencionados en la sección 5.2, ya que pueden capturar imágenes con gran resolución y gran campo visual de escenas 3D grandes y, además, permiten seleccionar a voluntad (dentro de ciertos límites) el paralaje. Estas características convierten a este tipo de sistema en el método de captura ideal para nuestra propuesta de generación de microimágenes. No obstante, para poder utilizar este sistema en nuestro método es necesario encontrar una forma rápida y efectiva de convertir las EIs capturadas en microimágenes sin que, a diferencia del algoritmo SPOC [75], se produzca una pérdida de información de la función plenóptica durante el proceso. Afortunadamente, realizar esto es posible si aprovechamos la equivalencia entre los sistemas de captura de InI y fotografía plenóptica que vimos en la sección 3.4.

Como ya comentamos, se puede considerar que las subimágenes calculadas a partir de una imagen plenótica son, en realidad, las EIs que capturaría un sistema de InI situado en el plano de la LC de la cámara plenótica [54], al que llamaremos sistema equivalente. Veamos si esta equivalencia también existe en sentido contrario, es decir, si es posible calcular microimágenes a partir de las EIs capturadas por un sistema de InI. Para ello, comenzamos calculando cómo sería la función plenótica que capturaría el sistema equivalente.

Consideremos la cámara plenótica de la Figura 3.15 (a), la función plenótica que registraría el sistema equivalente de la Figura 3.15 (b) se corresponde con la función plenótica en el plano justo antes de la refracción en la LC, de coordenadas (x', θ') . Para obtener la función plenótica en dicho plano, basta con aplicar las ecuaciones de transformación 2.3 y 2.5 sobre la función plenótica en el plano de la ML de la cámara plenótica, de coordenadas (x, θ) , de tal modo que:

$$\begin{pmatrix} x' \\ \theta' \end{pmatrix} = A^{-1} \begin{pmatrix} x \\ \theta \end{pmatrix}, \quad (5.1)$$

siendo A la matriz que define la transformación que sufre la función plenótica entre ambos planos y que se expresa como:

$$A = T_{a'} f_L = \begin{pmatrix} 1 - \frac{a'}{f_L} & -a' \\ 1/f_L & 1 \end{pmatrix}, \quad (5.2)$$

donde a' es la distancia desde la LC hasta el plano de la ML y f_L es la focal de la LC, Figura 3.15 (a). A partir de las ecuaciones 5.1 y 5.2 es inmediato calcular la función plenótica que obtendría el sistema equivalente:

$$\begin{pmatrix} x' \\ \theta' \end{pmatrix} = \begin{pmatrix} 1 & a' \\ -1/f_L & 1 - \frac{a'}{f_L} \end{pmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix} = \begin{pmatrix} x + a'\theta \\ -\frac{x + a'\theta - f_L\theta}{f_L} \end{pmatrix}. \quad (5.3)$$

Los diagramas espacio-angulares de la función plenóptica en los planos (x', θ') y (x, θ) se muestran en la Figura 5.8. Comparando ambos diagramas, podemos deducir dos aspectos de la función plenóptica que captura el sistema equivalente. Por un lado, existe una transposición entre la función plenóptica antes de la refracción en la LC y la correspondiente al plano de la ML, es decir, entre las microimágenes y las EIs del sistema equivalente. Por otro lado, como ya sucedía con la función plenóptica capturada por una cámara plenóptica, ver Figura 3.14, existe una cizalladura en la función plenóptica registrada por el sistema equivalente. Esto es debido a que, como vimos en la sección 3.4, las barreras que separan las distintas imágenes del sistema equivalente son oblicuas, o lo que es lo mismo, los ejes ópticos de las lentes o las cámaras del sistema equivalente se cruzan en un plano a una cierta distancia, a , que se corresponde con el PR de la cámara plenóptica, ver Figura 3.15 (b). Hay que destacar que el ángulo de la cizalla de la función plenóptica, ρ' , depende de la posición del PR de la cámara plenóptica:

$$\rho' = \frac{(f_L - a')}{f_L a'} = \frac{1}{a}, \quad (5.4)$$

donde se ha tenido en cuenta que, según la ecuación de conjugación de Gauss 3.1, $a' = f_L a / (a + f_L)$.

Ahora, consideremos el caso particular de una cámara plenóptica cuya ML se sitúa en la focal de la LC, $a' = f_L$, es decir, cuyo PR se encuentra en el infinito, $a = \infty$. En este caso, $\rho = 0$ y, por tanto, la función plenóptica que capturaría el sistema equivalente tomaría la forma que se muestra en la Figura 5.9. Podemos destacar dos características de dicha figura. Por un lado, la función plenóptica es similar a la que se obtendría con un sistema

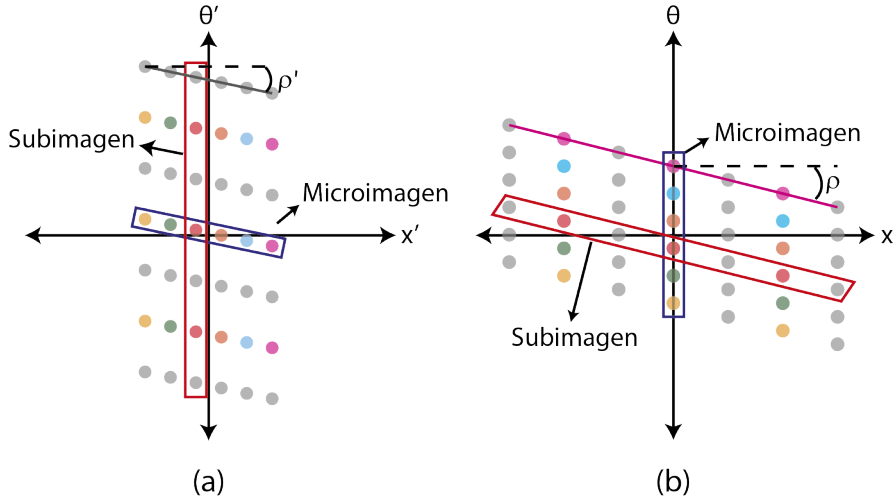


Figura 5.8.: (a) Función plenóptica antes de la refracción en la LC de una cámara plenóptica. (b) Función plenóptica en el plano de la ML.

de InI convencional, es decir, un sistema con los ejes ópticos de las lentes o cámaras paralelos entre sí, ver Figura 3.6. Por otro lado, la principal diferencia entre las Figuras 5.8 (a) y 5.9 es el valor de la cizalladura de la función plenóptica.

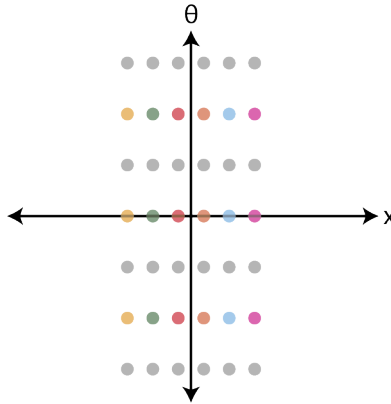


Figura 5.9.: Función plenóptica antes de la refracción en la LC de una cámara plenóptica en la que el PR se encuentra en el infinito, $a' = f_L$.

Teniendo en cuenta todo lo visto hasta el momento, podemos afirmar que existe una relación de equivalencia entre las EIs capturadas mediante

un sistema de InI y las EIs del sistema equivalente de una cámara plenóptica. Por tanto, concluimos que es posible que las EIs capturadas con una matriz de cámaras (cuyos ejes son paralelos entre sí) sean convertidas en microimágenes capturadas por un sistema de fotografía plenóptica. Para ello, debemos transponer filas y columnas de la función plenóptica codificada en las EIs. Podemos realizar la transposición usando el algoritmo para calcular subimágenes que vimos en la sección 3.3, ver Figura 3.12. Debido a la transposición, el número de microimágenes y píxeles por microimagen de la imagen plenóptica generada de esta manera será igual al número de píxeles por EI y al número de EIs, respectivamente. Asimismo, la posición del PR de la cámara plenóptica que habría capturado esta imagen puede ser controlada si, antes de realizar la transposición, se cizalla la función plenóptica codificada en las EIs. Nótese que de esta manera estaríamos controlando la posición del PR de la imagen plenóptica, es decir, estaríamos controlando la posición, con respecto a la MR, de la imagen 3D que se formaría al mostrar esta imagen plenóptica en un monitor de InI, ver sección 4.1.1.

El nuevo método de obtención de microimágenes que planteamos se basa en el proceso descrito en el párrafo anterior. Nuestra propuesta consiste en capturar una imagen integral para después convertirla en una imagen plenóptica mediante un algoritmo de transposición. Las principales ventajas de este método son, por un lado, que podemos controlar el número de microimágenes y el número de píxeles por microimagen de la imagen plenóptica generada, lo que permite adaptarla al monitor de InI en el que se va a mostrar. Por otro lado, con esta técnica podemos producir imágenes 3D ortoscópicas de escenas grandes y con gran paralelaje sin perder información de la función plenóptica capturada durante el proceso. Por último, también podemos controlar varios parámetros de la imagen 3D generada. Como este método presenta algunas similitudes con el algoritmo SPOC propuesto por Navarro y colaboradores [75], hemos decidido llamarlo SPOC 2.0 [65].

5.5 Resultados experimentales

Para comprobar el método propuesto se realizó el siguiente experimento. El primer paso consistió en obtener una imagen integral de una escena 3D compuesta por tres objetos: un tablón de madera que hace de fondo, una casa de Lego® y un muñeco, situados a 63 cm, 51 cm y 38 cm de la cámara, respectivamente. Para ello, en lugar de emplear una matriz de cámaras digitales, usamos el método de la apertura sintética³ [51]. La cámara digital empleada, una Canon 450 D, fue montada sobre una plataforma móvil controlada por un par de motores que se pueden desplazar en las direcciones (x, y) , y focalizada en el panel de madera que constituye el fondo. Todo el proceso de captura, desplazamiento, disparo y grabación de las fotos, fue controlado mediante un código de LabVIEW®. Los parámetros de la cámara durante la captura, $f = 18$ mm y $f/22$, se escogieron para que la profundidad de campo fuera lo suficientemente grande para obtener imágenes nítidas de toda la escena 3D.

La captura y preparación de la imagen integral se realizó de forma que las microimágenes obtenidas de ella se encuentren adaptadas al monitor de InI. Por tanto, teniendo en cuenta el número de píxeles por microimagen que el monitor de InI puede mostrar, obtuvimos un set de $N_H = N_V = 11$ EIs con un espaciado horizontal y vertical entre las cámaras de $P_H = P_V = 10$ mm, respectivamente. Como el campo visual capturado es mayor que la escena 3D de interés, aplicamos una máscara de recorte a cada EI para eliminar las partes exteriores, ver Figura 5.10. Nótese que mediante la selección del tamaño, λ , y la posición de la máscara de recorte, es posible modificar a voluntad el campo visual de las imágenes 3D mostradas. Tras el recorte, ajustamos el tamaño de cada EI a $n_H = n_V = 147$ píxeles para que, tal y como hicimos en la sección 5.1 con las subimágenes, fuera igual al número de microimágenes que

³Recordemos que en este método se capturan las distintas EIs con una única cámara digital que se desplazada mecánicamente.

el monitor de InI puede mostrar. En la Figura 5.11 mostramos la imagen integral obtenida.



Figura 5.10.: Ilustración del proceso de recorte. En este caso, la máscara de recorte se encuentra en el centro de todas las EIs.



Figura 5.11.: Imagen integral obtenida tras el proceso de captura, recorte y ajuste.

Para convertir esta imagen integral en microimágenes, se aplica el algoritmo de conversión propuesto: transponer filas y columnas en la función plenóptica. Para ello, basta con aplicar el algoritmo de cálculo de subimágenes, es decir, extraer los píxeles que ocupan la misma posición relativa en cada EI para componer una imagen. Este proceso proporciona una imagen plenóptica de $N_H = N_V = 147$ microimágenes de $n_H = n_V = 11$ píxeles cada una. No obstante, la imagen 3D que se genera en el caso de mostrar esta imagen plenóptica en el monitor de InI se

formará lejos del monitor y estará desenfocada. Por tanto, es necesario aplicar una cizalladura sobre la función plenóptica contenida en la imagen integral antes de realizar la transposición para, así, acercar la posición de la imagen 3D a los alrededores de la MR.

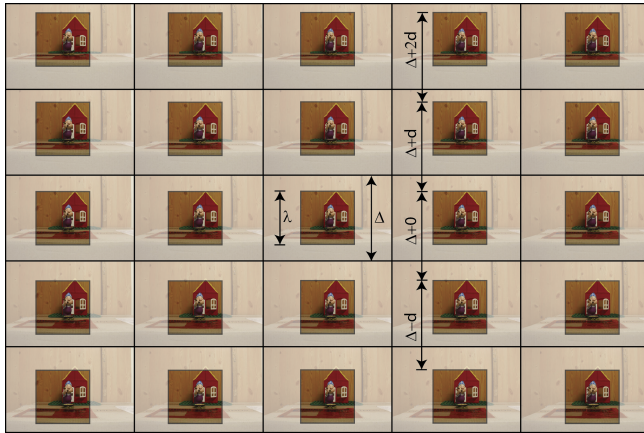


Figura 5.12.: Ilustración del desplazamiento de la máscara de recorte que provoca un cambio de la posición de la imagen 3D. En este caso, la posición de la máscara de recorte depende de la posición relativa de la EI dentro de la InI.

En lugar de cizallar la función plenóptica, podemos conseguir el efecto deseado si, cuando recortamos las EIs para eliminar las partes exteriores, desplazamos la máscara de recorte aplicada en función de la posición relativa de la EI dentro de la imagen integral. En concreto, este desplazamiento es proporcional a la distancia entre la correspondiente EI y la EI central, ver Figura 5.12. Como se puede deducir de la figura, la máscara de recorte que se aplica a cada EI se desplaza en un factor $i \cdot d$, donde i es un número entero que indica la posición de la EI siendo recortada con respecto a la EI central y d indica el desplazamiento proporcional, en píxeles, efectuado. Nótese que este proceso equivale a cambiar las barreras paralelas que separan las EIs durante la captura por barreras oblicuas, o lo que es lo mismo, equivale a cambiar el plano donde se cruzan los ejes ópticos de las cámaras empleadas en la captura, ver Figura 5.13.

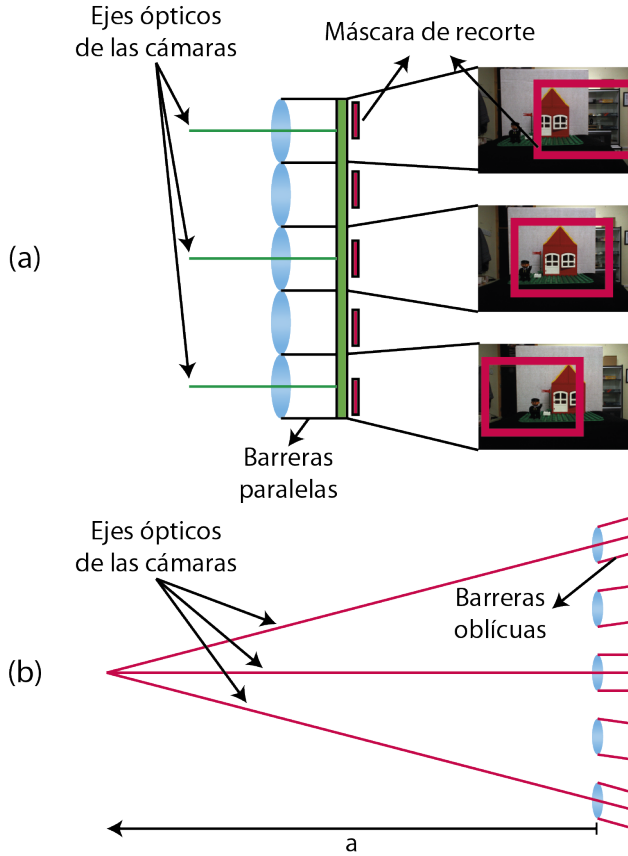


Figura 5.13.: (a) Representación esquemática de un sistema de captura de InI sobre cuyas Els se aplica una máscara de recorte siguiendo el método SPOC 2.0. (b) Sistema de captura equivalente que habría obtenido las Els recortadas en (a).

En la Figura 5.14 se muestra la relación entre dicho plano, a , (que será el PR de la imagen plenóptica) y el desplazamiento proporcional, en píxeles, de la posición del centro de recorte, d . Esta relación se puede expresar como:

$$d = \frac{(p_r g_r)}{a \delta}. \quad (5.5)$$

Si nos fijamos en la Figura 5.14 y la ecuación 5.5, podemos ver que d coincide con la disparidad⁴, en píxeles, que produciría un objeto, situado a una distancia a de la cámara, sobre dos EIs capturadas adyacentes.

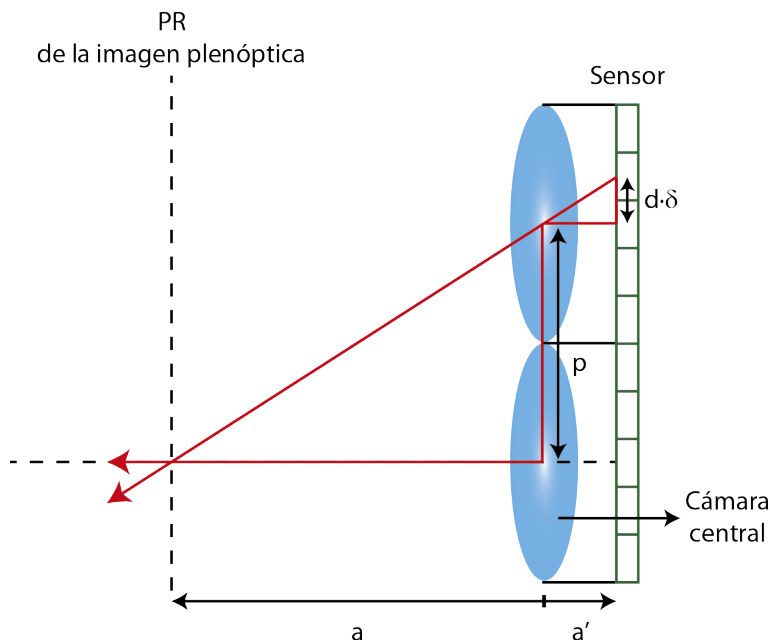


Figura 5.14.: Relación entre el desplazamiento en píxeles de la máscara de recorte y el punto en el que se cruzan los ejes ópticos de las cámaras del sistema de captura de InI. En la imagen, a' es la separación entre la lente de las cámaras y el sensor, p es el espaciado entre las cámaras y δ es el tamaño de los píxeles.

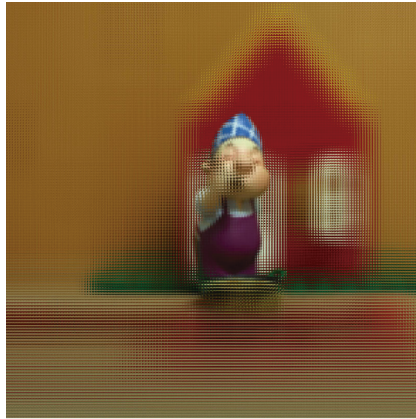
Utilizando este método, hemos creado tres conjuntos de microimágenes, ver Figura 5.15, preparados para ser mostrados en el monitor de InI que empleamos para generar la imagen 3D en la sección 5.1. En las Figuras 5.15 (a) y (b) el PR de la imagen plenóptica se fijó en la cabeza del muñeco. En la Figura 5.15 (c), el PR se fijó sobre la ventana de la casa. La diferencia entre las microimágenes que se muestran en las Figuras 5.15 (a) y (b) es el tamaño de la máscara de recorte aplicada, lo que provoca un cambio en el campo visual de la imagen 3D. Hay que tener

⁴Entendida como la diferencia en la posición aparente de la imagen de un objeto cuando es percibida desde dos puntos de vista diferentes. En este caso, la diferencia en la posición dentro de dos EIs.

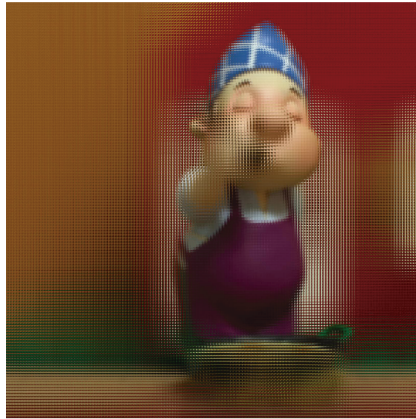
en cuenta que el tamaño en píxeles de las microimágenes del monitor de InI es un número no entero. Por tanto, antes de poder ser mostradas en él, es necesario ajustar ligeramente el tamaño en píxeles de las imágenes plenópticas de 1617×1617 a 1527×1527 para que el número de píxeles por microimagen sea el adecuado: $10,39 \times 10,39$.

En la Figura 5.16, se muestra el monitor de InI con los tres conjuntos de microimágenes generados. Adicionalmente, hemos preparado tres videos (**VÍDEO 5.2**, **VÍDEO 5.3** y **VÍDEO 5.4**) que simulan el proceso de visualización de la imagen 3D por parte de un observador monocular. Para ello, compusimos una película con las fotografías que hicimos del iPad mientras desplazábamos lateralmente una cámara en frente de él. En estos vídeos se puede comprobar que las imágenes 3D generadas son ortoscópicas y que tienen un gran paralaje. Aunque en las películas solo se muestra el paralaje horizontal, el monitor también presenta paralaje vertical.

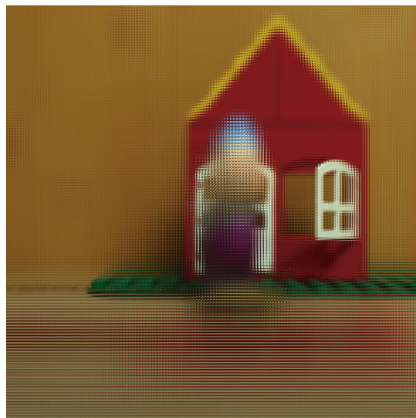
A partir de los resultados obtenidos en este experimento hemos comprobado que nuestro método permite aprovechar al máximo las características de un monitor de InI y que, además, permite generar imágenes 3D ortoscópicas con gran paralaje de escenas grandes y extensas. Adicionalmente, como en el proceso de captura se obtiene un excedente de campo visual y resolución espacial, es posible modificar ambos parámetros y la posición del PR de la imagen plenóptica a voluntad, lo que posibilita la creación de diferentes experiencias de visualización a partir de la misma imagen integral. Esta flexibilidad permite la rápida adaptación de las imágenes capturadas al monitor de InI empleado.



(a)



(b)

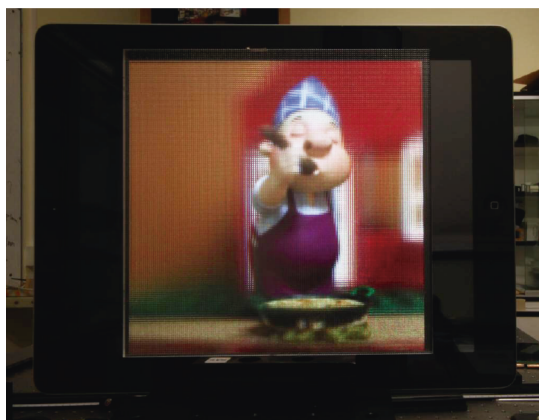


(c)

Figura 5.15.: Microimágenes calculadas a partir de la InI capturada.



(a)



(b)



(c)

Figura 5.16.: El monitor de InI mostrando los tres conjuntos de microimágenes generados.

Actualización del método SPOC 2.0

En el capítulo anterior propusimos un método de obtención de microimágenes, el SPOC 2.0 [65]. En el presente capítulo proponemos dos aplicaciones para este método que requieren actualizar ligeramente el algoritmo utilizado en SPOC 2.0. La primera de ellas consiste en aumentar el ángulo de visualización del monitor de InI gracias a incrementar el tamaño horizontal de las microimágenes [77]. La segunda aplicación consiste en la implementación de una rutina que permite generar videos 3D que simulen una cámara viajando en el interior de la escena, lo que se conoce como *trávelin* (del inglés *travelling*) [78].

6.1 Aumento del ángulo de visualización de un monitor de InI

Como vimos en la sección 4.2.1, un observador no siempre podrá percibir una correcta imagen 3D. Solo en el caso de que se encuentre en el interior de un AV, el observador verá una imagen 3D con paralaje completo. Uno de los principales inconvenientes de los monitores de InI es lo limitada que suele ser su área y ángulo de visualización. Debido a ello, se han realizado multitud de estudios intentando aumentar el ángulo de visualización [79]-[86].

En esta sección, proponemos utilizar el método SPOC 2.0 en conjunto con la propuesta realizada por Miura y colaboradores [84] para aumentar el ángulo de visualización de un monitor de InI [77].

6.1.1 Ángulo de visualización y rotación de la MR

Tanto el AV (ecuación 4.6) como el ángulo de visualización de un monitor de InI dependen del tamaño de las microimágenes que se muestran en la pantalla. El tamaño de estas microimágenes, en general, se corresponde con el tamaño de las lentes de la MR, p_r . El ángulo de visualización horizontal, α , de un monitor de InI se muestra en la Figura 6.1.

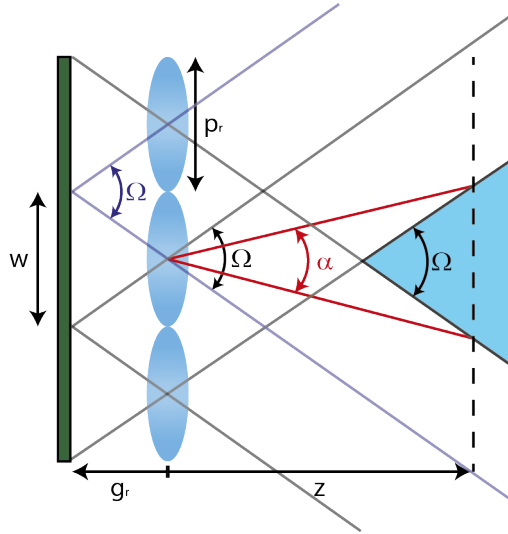


Figura 6.1.: Ángulo de visualización horizontal de un monitor de InI. Ω representa el ángulo de visualización de la microimagen central.

A partir de esta imagen podemos deducir fácilmente la expresión para el ángulo de visualización horizontal, α , y vertical, β :

$$\alpha = 2 \tan^{-1} \left(\frac{w}{2g_r} - \frac{w(l-1)}{2z} \right), \quad (6.1)$$

$$\beta = 2 \tan^{-1} \left(\frac{h}{2g_r} - \frac{h(l-1)}{2z} \right), \quad (6.2)$$

donde w representa la anchura de la microimagen, h representa su altura, z representa la distancia desde el observador a la MR y l es el número de imágenes que se muestran en el monitor de InI. El primer termino del lado derecho de las ecuaciones 6.1 y 6.2, representa el ángulo de visualización de la microimagen central, Ω .

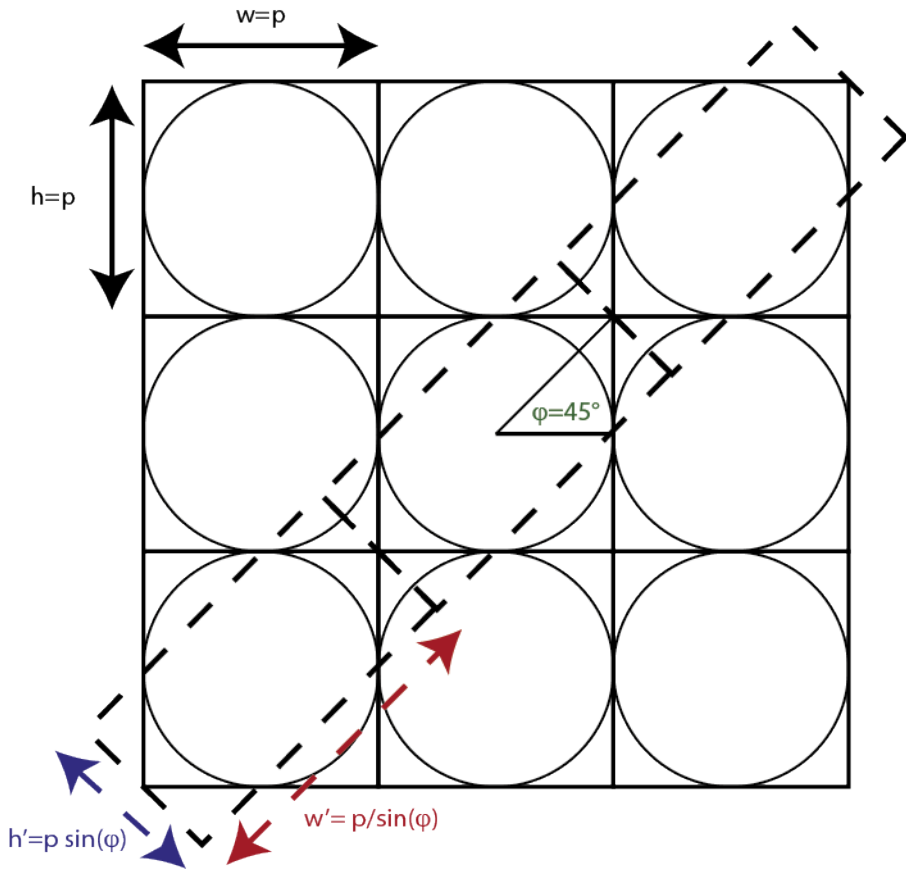


Figura 6.2.: Comparación del tamaño y distribución de las microimágenes de una matriz de lentes de distribución rectangular con y sin rotación. Las líneas continuas son las microimágenes que se capturan de manera convencional. Las líneas a rayas son las microimágenes capturadas con la misma matriz pero girada un ángulo $\varphi = 45^\circ$.

Para aumentar el ángulo de visualización horizontal, Miura y colaboradores [84] propusieron modificar el tamaño de las microimágenes para que éste fuera diferente al tamaño de las lentes del monitor de InI. En su propuesta, esta modificación debía realizarse manteniendo invariante el número total de píxeles de las microimágenes, de tal forma que al incrementar la anchura de éstas, se redujera su altura. Para ello, rotaron la matriz de lentes durante el proceso de captura y reconstrucción de la imagen 3D. En la Figura 6.2, se observa el cambio que sufren las microimágenes cuando se rota una matriz de lentes de distribución rectangular. El tamaño de éstas microimágenes no convencionales, a las que llamaremos microimágenes aumentadas, puede ser calculado con la ecuación 6.3:

$$w' = \frac{p}{\sin \varphi} \quad y \quad h' = p \sin \varphi. \quad (6.3)$$

Los ángulos de visualización del monitor cuando se muestran microimágenes aumentadas se expresan de la siguiente forma:

$$\alpha' = 2 \tan^{-1} \left(\frac{p}{2 g_r \sin \varphi} - \frac{p(l-1)}{2 z \sin \varphi} \right), \quad (6.4)$$

$$\beta' = 2 \tan^{-1} \left(\frac{p \sin \varphi}{2 g_r} - \frac{p \sin \varphi (l-1)}{2 z} \right). \quad (6.5)$$

En las ecuaciones anteriores, φ representa el ángulo que se ha rotado la matriz de lentes. Nótese que el ángulo de visualización horizontal aumenta a consta de reducir el ángulo de visualización vertical.

Aunque el método propuesto por Miura y colaboradores es una forma sencilla y buena de mejorar el ángulo de visualización, tiene la importante desventaja de que es muy limitado: sólo funciona en el caso de que los sistemas de captura y reconstrucción trabajen en modo simétrico.

6.1.2 Combinación con SPOC 2.0

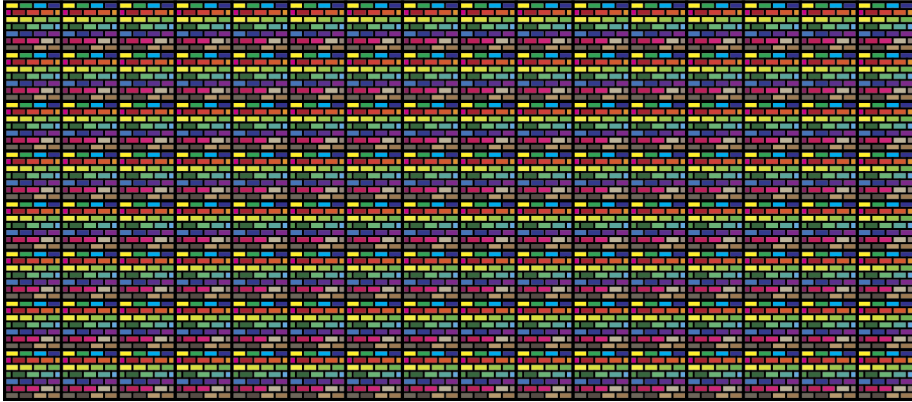
Para aumentar la utilidad de la idea de Miura y colaboradores [84], es necesario conseguir que ésta también puede ser implementada en monitores de InI que no sean simétricos con respecto al sistema de captura. Para ello, proponemos usar el método SPOC 2.0 [65] para generar una imagen plenóptica compuesta por microimágenes aumentadas, es decir, microimágenes cuyo tamaño y distribución se corresponde con el de una MR girada un determinado ángulo φ , ver Figura 6.2. No obstante, será necesario modificar el algoritmo utilizado en SPOC 2.0 para que las microimágenes obtenidas tengan una distribución ajustada a una MR que se encuentra girada un ángulo φ .

Como vimos en el capítulo anterior, cuando generamos una imagen plenóptica con el método SPOC 2.0, las características de ésta vienen determinadas por la imagen integral capturada, ver sección 5.4.1. Por ejemplo, el tamaño y el número de las microimágenes que se generan depende, respectivamente, del número y del tamaño de las EIs capturadas. En consecuencia, es posible que la distribución espacial de los píxeles de las microimágenes también venga determinada por alguna característica de las EIs. Si recordamos el algoritmo de conversión entre EIs y microimágenes, cada microimagen de la imagen plenóptica se forma extrayendo el mismo pixel de cada EI, de tal modo que el pixel j de cada una de las EIs constituye la microimagen j de la imagen plenóptica. Asimismo, la posición dentro de la microimagen del pixel extraído de cada EI depende de la posición relativa de ésta dentro de la imagen integral. Entonces, podemos asumir que cada pixel en la EI se relaciona con una microimagen y, por lo tanto, la posición relativa de los píxeles de una EI determina la distribución de las microimágenes en la imagen plenóptica. Por ejemplo, en una cámara digital convencional, los píxeles se encuentran distribuidos en una rejilla rectangular y, por tanto, la distribución de las microimágenes que se generan con SPOC 2.0 será equivalente. Sin embargo, si los píxeles de las EIs estuvieran distribuidos en una rejilla hexagonal, las

microimágenes generadas se encontrarían distribuidas también siguiendo un patrón hexagonal. Por tanto, para conseguir que el método de SPOC 2.0 genere imágenes plenópticas con una distribución determinada de microimágenes necesitamos modificar la distribución de los píxeles de las EIs capturadas. Esto puede conseguirse fácilmente si se emplean cámaras digitales para obtener las EIs. Como las EIs capturadas con una cámara digital tienen una cantidad de píxeles mayor a la necesaria, se pueden promediar o combinar sus píxeles para conseguir la distribución deseada, ver Figura 6.3.

Teniendo en cuenta lo comentado en el párrafo anterior, procedemos a describir los pasos necesarios para producir una imagen plenóptica compuesta por microimágenes aumentadas, a la cual llamaremos imagen plenóptica aumentada, usando el método de SPOC 2.0 actualizado. Primero, se debe elegir el monitor de InI que se va a utilizar para producir la imagen 3D. Segundo, se calcula el número de píxeles que tendrían las microimágenes del monitor de InI elegido en el caso convencional, es decir, cuando la MR no se encuentra rotada. Tercero, decidimos la rotación que tendrá la MR, lo que determina el incremento en el ángulo de visualización. Nótese que, debido a la superposición de las microimágenes, no todos los ángulos de rotación son válidos [83]. Cuarto, calculamos el nuevo tamaño en píxeles, usando las ecuaciones 6.4 y 6.5 en el caso de utilizar una MR de distribución rectangular, de las microimágenes aumentadas, es decir, las que se mostrarán con la MR rotada. Quinto, se captura una imagen integral con un número de EIs igual al número de píxeles por microimagen de la imagen plenóptica aumentada. Finalmente, transformamos la imagen integral en su correspondiente imagen plenóptica aplicando el algoritmo de transposición de SPOC 2.0, pero cambiando el número de píxeles de las EIs para que tengan la misma distribución que las lentes de la MR rotada. Nótese que en el caso de que el número de píxeles por microimagen aumentada sea no entero, es necesario ajustar el tamaño de la imagen plenóptica para que tenga el número correcto de píxeles por microimagen.

Els capturadas con una cámara



Microimágenes obtenidas de forma computacional

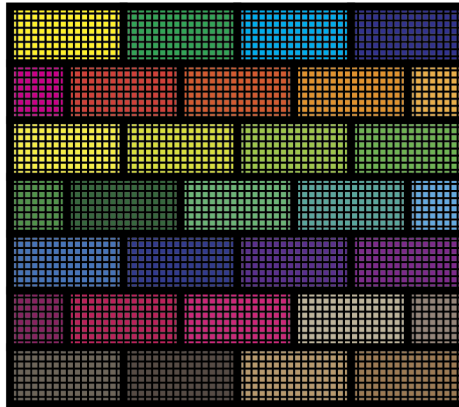


Figura 6.3.: (a) Ejemplo de combinación de píxeles que produce que los píxeles de una imagen integral tengan una distribución diferente. (b) Microimágenes generadas a partir de la imagen integral mostrada en (a).

La imagen obtenida con el procedimiento descrito en el párrafo anterior puede ser mostrada en el monitor de InI con la MR rotada y generar una imagen 3D que, con respecto al caso convencional, tiene un mayor ángulo de visualización horizontal y menor ángulo de visualización vertical. Hay que destacar que no solo se produce un cambio en el ángulo de visualización, si no que, además, también cambia el paralaje (o número de vistas) que se muestran de la imagen 3D.

6.1.3 Resultados experimentales

Para comprobar la validez de nuestra propuesta realizamos un experimento en el que capturamos dos imágenes integrales, las transformamos en imágenes plenópticas y, finalmente, las mostramos en un monitor de InI. Mientras que una de las imágenes plenópticas se obtuvo con el método SPOC 2.0 convencional, con microimágenes cuadradas, la otra se generó para que tuviera el ángulo de visualización horizontal aumentado, con microimágenes aumentadas. Ambas capturas fueron realizadas desplazando una única cámara tras cada fotografía utilizando la técnica de la apertura sintética [51].

Como monitor de InI reutilizamos el empleado en el capítulo anterior. Éste está constituido por un iPad con una densidad de píxeles de 10,39 píxeles/mm, que hace de pantalla, y una MR de distribución rectangular con un tamaño y focal de $p_r = 1$ mm y $f_r = 3,3$ mm, respectivamente. En consecuencia, las microimágenes de la imagen plenóptica convencional tienen un tamaño de $w = 1$ mm y $h = 1$ mm, el mismo que el de las lentes de la MR. Por otro lado, en nuestro experimento el ángulo de rotación de la MR es de $\varphi = 45^\circ$, lo que establece el tamaño de las microimágenes aumentadas en $w = \sqrt{2}$ mm y $h = 1/\sqrt{2}$ mm.

Para capturar las imágenes integrales, utilizamos una cámara Canon 450D con un objetivo de distancia focal 18 mm, enfocado al infinito. La resolución de las fotografías tomadas es de 4272×2848 píxeles, con una

distancia entre fotos de $P_H = P_V = 5$ mm. Al igual que en el capítulo anterior, la posición de la cámara se controla con un par de motores que pueden desplazarse en las direcciones (x, y) . La escena 3D capturada está compuesta por un fondo, un juguete y dos cilindros de acero situados a 52,5 cm, 40 cm y 35 cm de la cámara, respectivamente.

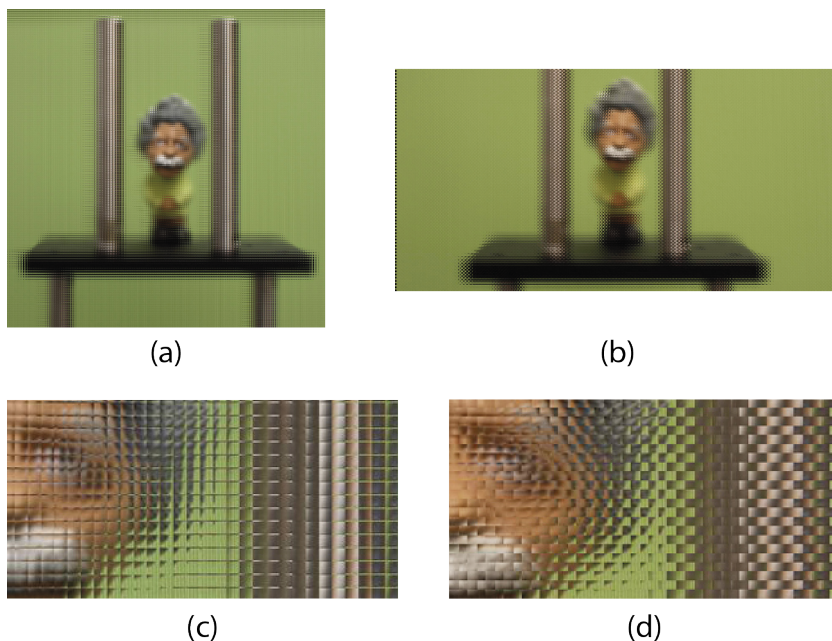


Figura 6.4.: (a) Imagen plenóptica convencional con 139×139 microimágenes y $10,39 \times 10,39$ píxeles por microimagen, (b) imagen plenóptica aumentada con 139×139 microimágenes de $14,69 \times 7,35$ píxeles, (c) microimágenes de la imagen plenóptica convencional, y (d) microimágenes de la imagen plenóptica aumentada.

Para obtener las microimágenes correctas para cada una de las imágenes plenópticas, el número de EIs de la imagen integral capturada debe ser igual al número de píxeles por microimagen. Sin embargo, como solo se puede capturar un número entero de fotos, la cantidad de fotos realizada para el caso de la imagen convencional es 11×11 , mientras que para la imagen aumentada es 15×8 . Sobre estas imágenes integrales aplicamos el algoritmo de SPOC 2.0 para obtener dos imágenes plenópticas de 139×139 microimágenes. Después de la transformación, ajustamos el tamaño de

las microimágenes para que tengan el número correcto de píxeles. Las imágenes finales se muestran en la Figura 6.4.

Para comprobar el aumento en el ángulo de visualización, medimos el cambio sufrido en el ángulo de visualización de la microimagen central, Ω . Para ello, realizamos una serie de fotografías simulando el movimiento de la cabeza de un observador que esté visualizando las imágenes 3D generadas. La distancia entre fotografías se corresponde con una diferencia angular de $0,5^\circ$. Si nos fijamos en la Figura 6.1, el observador habrá recorrido un ángulo Ω cuando perciba que la perspectiva de la imagen 3D se repite, lo que nos proporciona unos ángulos de visualización de $15^\circ \pm 1^\circ$ para el caso convencional y $22^\circ \pm 1^\circ$ para la imagen aumentada, aproximadamente, ver Figuras 6.5 y 6.6.

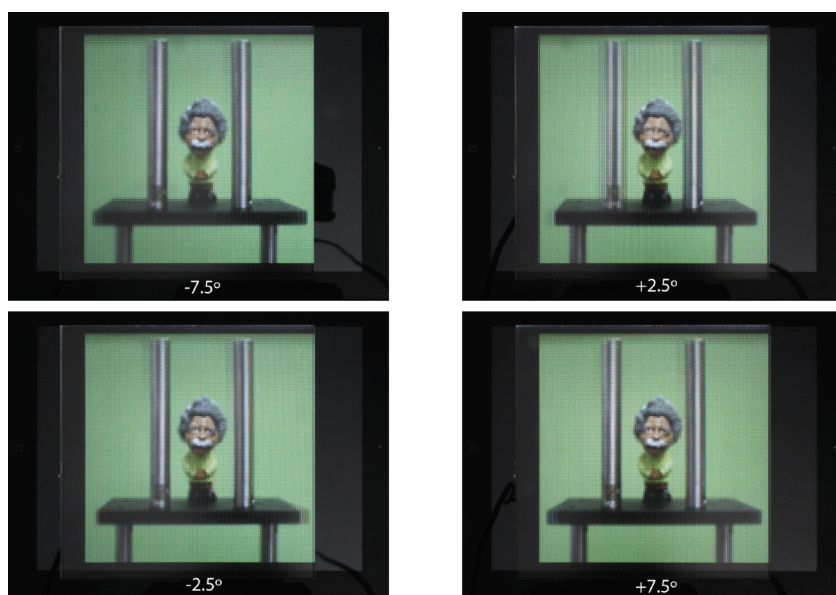


Figura 6.5.: Selección de fotografías que se corresponden con un periodo completo del ángulo de visualización para la imagen convencional.

Aunque queda claro que el el ángulo de visualización ha aumentado, nos interesa comparar el valor experimental con los valores teóricos. Si tenemos en cuenta las características del monitor de InI implementado (ver anexo A) obtenemos unos valores teóricos de $15,0^\circ$ y $21,1^\circ$ para el

caso convencional y aumentado, respectivamente. Vemos que los valores teóricos obtenidos se encuentran dentro del error experimental.

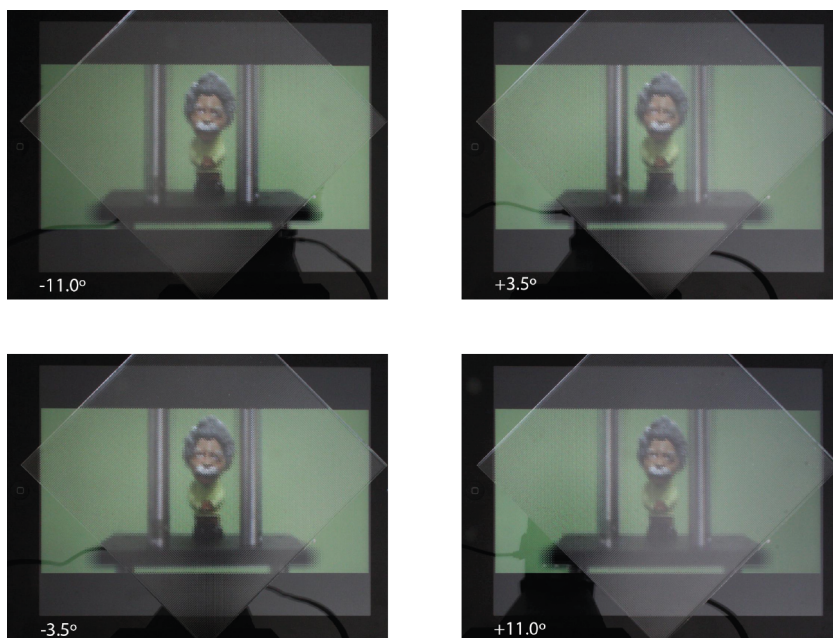


Figura 6.6.: Selección de fotografías que se corresponden con un periodo completo del ángulo de visualización para la imagen aumentada.

6.2 Película 3D a partir de una única imagen integral

Si en la sección 6.1 proponíamos utilizar una versión mejorada del método SPOC 2.0 para incrementar el ángulo de visualización de un monitor de InI, ahora exponemos una aplicación diferente de este método dirigida a la producción de contenido visualizable en 3D. Específicamente, proponemos una técnica para la generación de películas 3D en las que se muestra un trávelin de la escena a partir de una única imagen integral [78]. Esto es posible gracias a que el método de SPOC 2.0 permite el control del campo visual y el PR de la imagen plenóptica generada.

6.2.1 Generar secuencias de imágenes con SPOC 2.0

Como vimos en el capítulo anterior, el campo visual y el PR de la imagen plenóptica generada con SPOC 2.0 se ajustan controlando la posición y el tamaño de las máscaras de recorte aplicadas sobre las EIs, ver sección 5.5. De este modo, cada vez que se aplica el algoritmo de SPOC 2.0 sobre la misma imagen integral, pero cambiando los parámetros para la máscara de recorte, se genera una colección diferente de microimágenes, ver Figura 6.7.

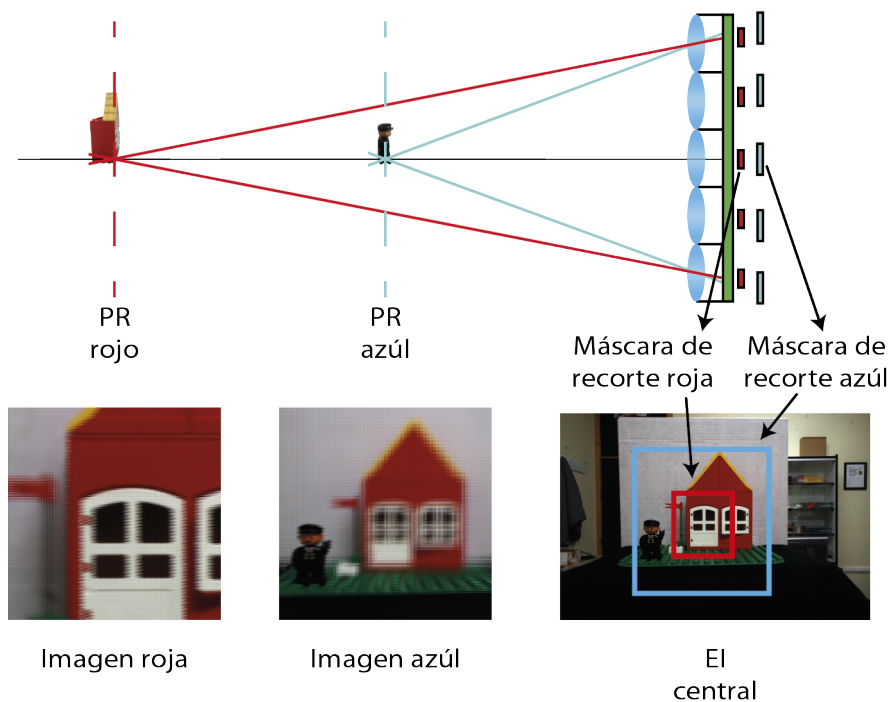


Figura 6.7.: Esquema que muestra como el cambio del tamaño y la posición de la máscara de recorte que se aplica sobre cada EI modifica el campo visual y la posición del PR de la imagen plenóptica generada. En concreto, se aplican dos máscaras de recorte distintas, roja y azul, sobre la misma imagen integral.

A partir de lo comentado en el párrafo anterior, podemos concluir que es posible utilizar el método de SPOC 2.0 para generar una secuencia de imágenes plenópticas que constituyan los fotogramas de un vídeo 3D que muestre un tróvelin de la escena 3D capturada. Para ello, debemos aplicar varias veces el algoritmo de dicho método sobre la imagen integral capturada, cambiando cada vez los parámetros de la máscara de recorte para conseguir el efecto deseado.

6.2.2 Mapa de disparidad

Para facilitar la generación de la secuencia de imágenes, hemos añadido una actualización al algoritmo de SPOC 2.0 que ayuda a situar con precisión la posición del PR en cada una de las imágenes plenópticas generadas.



Figura 6.8.: Mapa de disparidad obtenido a partir de dos EIs utilizando el método SAD. Cada color representa un valor de disparidad distinto.

Cuando explicamos el proceso de recorte del algoritmo de SPOC 2.0 en la sección 5.5, comentamos que el desplazamiento proporcional que se aplicaba a la máscara de recorte, d , se correspondía con el valor de la disparidad que produciría un objeto, situado a una distancia a de las

cámaras, sobre dos EIs capturadas adyacentes. Esto implica que para conocer el valor de d necesario para situar el PR de las imágenes plenópticas generadas con SPOC 2.0 sobre determinados objetos de la escena 3D, basta con conocer la disparidad que producen en el sistema de captura. Por ello, hemos incorporado una rutina al algoritmo de SPOC 2.0 que obtiene un mapa de disparidad a partir de dos de las EIs obtenidas en el proceso de captura. Como ejemplo, en la Figura 6.8 mostramos un mapa de disparidad de la escena 3D que se muestra en la Figura 6.9, y que ha sido calculado con el método SAD (del inglés *Sum of Absolute Differences*). Nótese que a partir de este mapa y la ecuación 5.5, es posible calcular la posición (x, y, z) de los objetos de la escena 3D.

6.2.3 Resultados experimentales.

Para ilustrar nuestro método generamos tres películas, obtenidas de tres experimentos distintos, y las mostramos en un monitor de InI [78].

Como monitor de InI volvemos a utilizar el mismo que en el capítulo anterior y la sección 6.1.3: un iPad como pantalla y una MR con lentes cuyos tamaños y distancias focales son, respectivamente, $p_r = 1$ mm y $f_r = 3,3$ mm. Por tanto, para que las películas se reproduzcan adecuadamente en este monitor de InI, sus fotogramas deberán estar compuestos por imágenes plenópticas con 147×147 microimágenes de $10,39 \times 10,39$ píxeles cada una. Esto implica que las imágenes plenópticas generadas con SPOC 2.0 deberán ser ajustadas como hicimos en la sección 5.5.

Para el primer experimento capturamos, con la técnica de la apertura sintética [51], una colección de 11×11 EIs separadas $P_H = P_V = 6$ mm entre sí. La escena 3D utilizada en este caso se muestra en la Figura 6.9. La distancia entre la cámara y el primer objeto es de 39 cm, siendo la profundidad total de la escena, desde el muñeco hasta el fondo, de unos 16 cm. Para realizar la captura, utilizamos una cámara Canon 450D con un objetivo de distancia focal 18 mm, enfocado al infinito. En la Figura 6.10

mostramos un subconjunto de la imagen integral capturada compuesto por 5×5 EIs.



Figura 6.9.: Escena 3D usada en nuestro primer experimento.



Figura 6.10.: Conjunto de 5×5 EIs, de la escena 3D mostrada en 6.9, obtenidas con la técnica de la apertura sintética.

Para generar las imágenes plenópticas que compondrán los fotogramas de la película 3D, aplicamos múltiples veces el algoritmo de SPOC 2.0 sobre las EIs capturadas. Cada vez que dicho algoritmo fue aplicado, el proceso de recorte de las EIs fue diferente. Con ello, conseguimos que el campo visual y el PR de la secuencia de imágenes resultante simule el movimiento de una cámara por el interior de la escena 3D. Como ejemplo, en la Figura 6.11 mostramos dos conjuntos de EIs recortadas. Las EIs recortadas que se muestran en la Figura 6.11 (a) se calcularon para que el campo visual y el PR de la imagen plenóptica generada a partir de ellas estuviera centrado en la bandera. Por otro lado, las EIs de la Figura 6.11 (b) proporcionaron una imagen plenóptica con el campo visual y el PR centrados en la puerta.

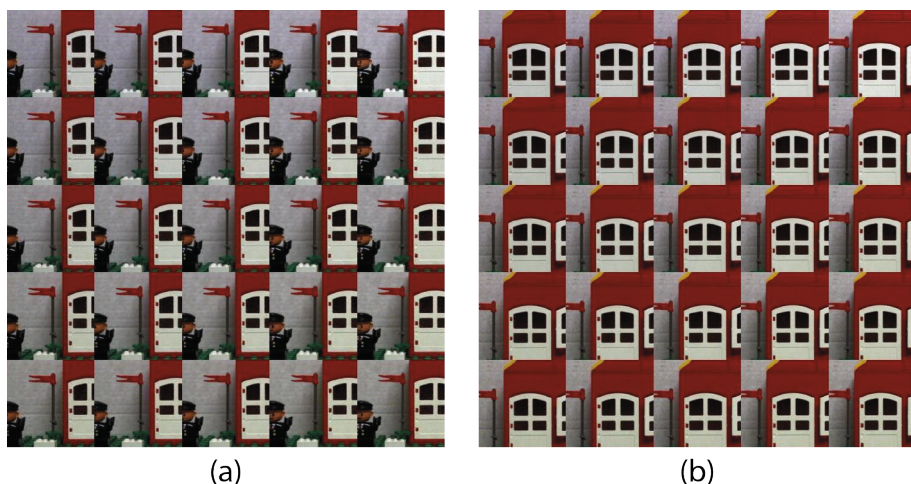


Figura 6.11.: EIs sintéticas obtenidas tras el proceso de recorte. (a) Campo visual y PR centrados en la bandera. (b) Campo visual y PR centrados en la puerta.

El número de imágenes plenópticas que calculamos en total fue 126. Con ellas, compusimos la película, **VÍDEO 6.1**. Dos fotogramas de este vídeo se muestran en la Figura 6.12. Hay que destacar que cuando esta película se reproduce en el monitor de InI, las partes de la escena que se reconstruyen delante, sobre o detrás de la MR cambian conforme cambia el PR en cada fotograma de la película.

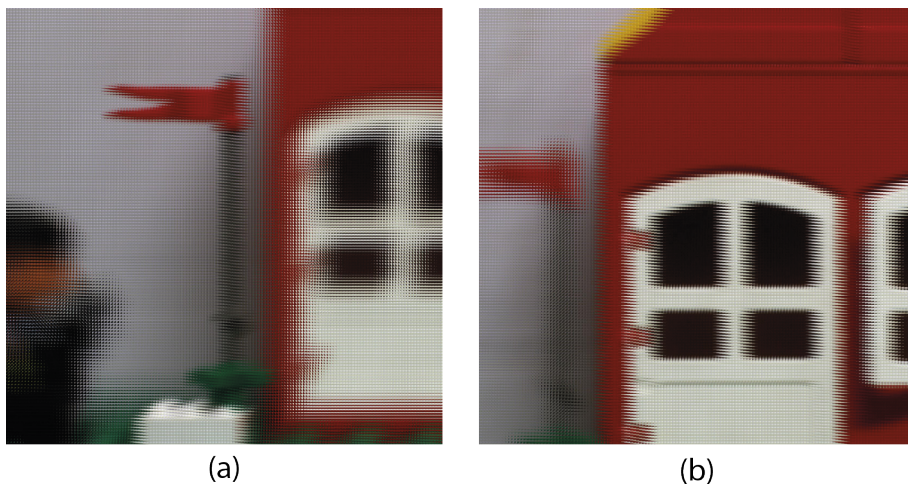


Figura 6.12.: Imágenes plenópticas calculadas a partir de la imagen integral mostrada en la Figura 6.10.

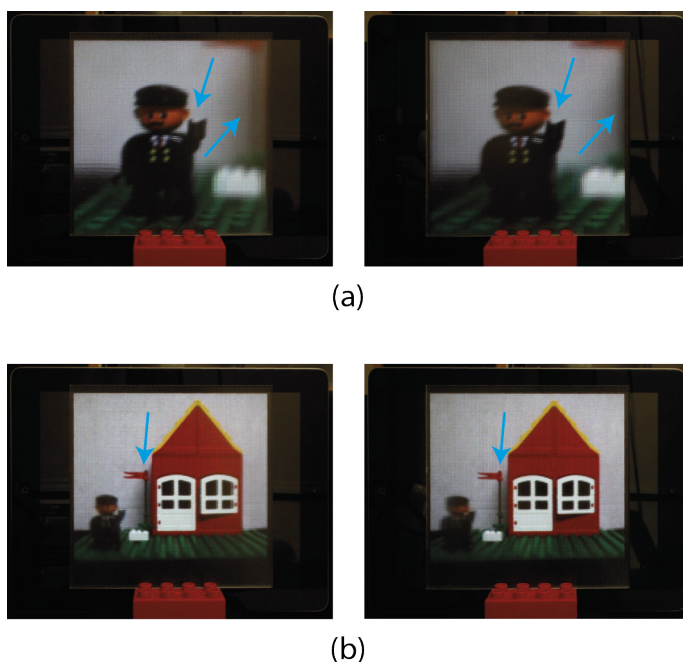


Figura 6.13.: Fotogramas de la película **VÍDEO 6.2**. Tanto en (a) como en (b) se muestra el momento en el que el monitor de InI reproduce el mismo fotograma de **VÍDEO 6.1** pero en dos posiciones distintas de la cámara.

Para terminar este primer experimento, simulamos el proceso de visualización de esta primera secuencia 3D por parte de un observador monocular. Para ello, grabamos un nuevo vídeo con una cámara situada a una cierta distancia del monitor, a 50 cm aproximadamente, y que se desplazaba en frente del monitor de InI para percibir la diferencia en paralaje. Esta película se muestra en **VÍDEO 6.2**. Como se puede apreciar en dicho vídeo, el trávelin a través de la escena se detiene en tres momentos. Esto se hizo para que la cámara pudiera percibir correctamente el paralaje de la imagen 3D. En la Figura 6.13 se muestran cuatro fotogramas extraídos de este vídeo a los que hemos añadido flechas señalando las diferencias en paralaje. En la imagen (a) de la Figura 6.13, se puede apreciar que el campo visual y el PR de la imagen reconstruida se encuentran centrados en el muñeco. Por otro lado, en la imagen (b) el campo visual abarca la escena 3D al completo, mientras que el PR está situado en el plano que contiene la casa. Asimismo, en ambos casos se muestran dos vistas distintas de la escena que se corresponden con dos posiciones distintas de la cámara.

Realizamos un segundo experimento con el objetivo de comprobar si es posible aplicar el efecto trávelin sobre un vídeo de InI. Para ello, capturamos un conjunto de doce imágenes integrales en las cuales la posición de un muñeco iba cambiando. Para reducir el tiempo y esfuerzo necesario para realizar la captura de este vídeo de InI, el número de EIs capturadas por imagen integral fue de 5×5 . Decidimos que el trávelin siguiera la trayectoria del muñeco, por tanto, el campo visual y el PR de las imágenes generadas siempre se encontrarán centrados en el plano donde se encuentra éste. Hay que destacar que, en este caso, sólo se genera una imagen plenóptica a partir de cada imagen integral. Tras aplicar todo el proceso obtenemos un vídeo compuesto de doce fotogramas, ver **VÍDEO 6.3**. Nótese que para poder mostrar los fotogramas en el monitor de InI fue necesario aumentar el número de píxeles por microimagen de 5×5 a $10,39 \times 10,39$. En consecuencia, como comentamos en la sección 5.1, la calidad de las imágenes plenópticas será inferior a las obtenidas con el

número correcto de píxeles por microimagen. Al igual que hicimos en el primer experimento, grabamos un vídeo con una cámara situada en frente del monitor de InI mientras este último reproduce **VÍDEO 6.3**. Cuatro de los fotogramas de esta película, **VÍDEO 6.4**, se muestran en la Figura 6.14.

Para corroborar los resultados anteriores se repitió el experimento con otra escena 3D y un número de fotos distinto. La película obtenida se muestra en **VÍDEO 6.5**.



Figura 6.14.: Fotogramas de la película generada en el experimento de visualización de **VÍDEO 6.3**.

Mejora de la experiencia de visualización de un monitor de InI mediante un dispositivo eye-tracking

En los últimos capítulos hemos propuesto y mejorado un método, el SPOC 2.0 [65], que nos permite obtener microimágenes y adecuarlas a las características del monitor de InI en el que se van a mostrar.

En este capítulo, planteamos un nuevo método para incrementar el AV de un monitor de InI mediante el uso de un sistema que sigue el movimiento de los ojos (del inglés *eye-tracking*) y el método SPOC 2.0 [87]. Adicionalmente, con esta propuesta aumentamos el paralaje que el monitor de InI puede mostrar y, también, permitimos al observador elegir qué partes de la imagen 3D se reconstruyen delante o detrás de la MR. Este método se basa en la modificación en tiempo real y en función de la posición 3D del observador de las microimágenes mostradas en el monitor de InI. Para seguir los movimientos del observador utilizamos un dispositivo Kinect, mientras que los cambios en la información mostrada en el monitor se realizan con el algoritmo de SPOC 2.0.

7.1 Aumento del área de visualización y paralaje mediante un dispositivo eye-tracking y el método SPOC 2.0.

La utilización de un sistema de eye-tracking para aumentar el ángulo de visualización de un monitor de InI no es algo nuevo. Los primeros en usarlo fueron Park y colaboradores [88], [89], que determinaron la posición del observador con la ayuda de una cámara infrarroja (IR) y dos LEDs IR instalados en unas gafas. Con esta información, pudieron adaptar las imágenes mostradas en el monitor de InI a la posición del observador. Otra aproximación a esta idea fue planteada por Hong y colaboradores [90], quienes usaron un dispositivo Kinect para seguir la posición del observador. Más adelante, Shen y colaboradores [91] volvieron a utilizar un dispositivo de seguimiento compuesto por una cámara IR y un LED. La principal aportación de su propuesta fue la adaptación de las microimágenes a una posible rotación de la cabeza del observador. Finalmente, Xiong y colaboradores [92] utilizaron un enfoque diferente ya que, para el cálculo computacional de las microimágenes usaron una matriz de cámaras virtuales ortográficas con el mismo punto de convergencia. No obstante, estas propuestas presentan dos inconvenientes importantes. El primero de ellos es que sólo han sido demostradas para imágenes 3D generadas por ordenador, mientras que el segundo inconveniente es debido a que las escenas 3D que fueron utilizadas son muy sencillas ya que, en general, estaban compuestas por dos objetos planos y siempre sobre fondo negro.

7.1.1 Método

A continuación, proponemos un nuevo método para aumentar el AV de un monitor de InI y el paralaje que éste puede mostrar de la escena 3D mediante el uso de un dispositivo eye-tracking. No obstante, a diferencia

de otras propuestas anteriores, en nuestro caso utilizamos un sistema de captura y escena 3D reales. Asimismo, con nuestro método también podemos adaptar el PR de la imagen 3D a la posición en profundidad del observador.

El primer paso de nuestra propuesta es la captura de una imagen integral de una escena 3D real. Para que nuestro método funcione, debemos capturar un exceso de información angular. Es decir, en lugar de realizar la captura de la imagen integral para que las microimágenes calculadas a partir de ella se encuentren adaptadas al monitor de InI, como hemos hecho hasta el momento en esta Tesis, capturamos un exceso de vistas y paralaje. De este modo, las microimágenes que se obtengan al aplicar el algoritmo SPOC 2.0 serán mayores que las lentes de la MR y, por tanto, deberán ser adaptadas. No obstante, en lugar de ajustar su tamaño interpolando, como propusimos en la sección 5.1, las recortaremos. Tener un exceso de perspectivas permite seleccionar qué vistas compondrán las microimágenes recortadas y, por tanto, controlar cuales serán mostradas en el monitor y percibidas por el observador. Tras el proceso de recorte, que explicaremos en detalle más adelante, el siguiente paso de nuestro método consiste en cambiar dinámicamente las microimágenes que se muestran en el monitor de InI en función de la posición del observador. Para ello, utilizamos un sistema de eye-tracking, en nuestra propuesta un dispositivo Kinect, que proporciona a un software de control la posición 3D del observador en tiempo real. Con dicha información, el software de control puede modificar las microimágenes mostradas dinámicamente. En concreto, el programa desplaza las microimágenes que se muestran en el monitor, cambia sus perspectivas y cambia el PR. A continuación detallamos estos procesos y qué efectos tienen sobre la experiencia de visualización del observador.

Proceso de recorte y desplazamiento de las microimágenes

Tras la conversión de las EIs capturadas en microimágenes con SPOC 2.0, tenemos a nuestra disposición un conjunto de microimágenes que son significativamente más grandes que las lentes del monitor de InI en el que se van a mostrar. Para ajustar las microimágenes al tamaño correcto, aplicamos sobre cada una de ellas una máscara de recorte, siendo el tamaño de la máscara igual al de las lentes de la MR. La posición de la máscara de recorte y la posición de las microimágenes recortadas que se muestran en el monitor de InI son controladas por el software de control y dependerán de la posición lateral del observador. Cuando este último se encuentra justo delante de la lente central de la MR, el software de control colocará la máscara de recorte en el centro de las microimágenes y mostrará las microimágenes recortadas de manera convencional, ver Figura 7.1 (a). Sin embargo, cuando el observador se mueva lateralmente, las máscaras serán desplazadas proporcionalmente, cambiando así las vistas de las microimágenes resultantes. Asimismo, estas microimágenes no se mostrarán alineadas con respecto a sus correspondientes lentes. En su lugar, serán desplazadas sobre la pantalla del monitor en la misma proporción para que el AV siga el movimiento del observador, ver Figura 7.1 (b). Como el AV se desplaza con el observador, el ángulo de visualización, α , aumenta considerablemente en comparación con la configuración convencional.

En nuestra propuesta, el desplazamiento del observador necesario para que el software de control realice estos cambios coincide con la distancia que el observador debe moverse lateralmente para cambiar la vista de la imagen 3D que percibe. Como éste solo ve un píxel tras cada lente, véase sección 4.2.2, esto sucederá cuando el observador pase a ver un píxel adyacente. Esta distancia puede ser calculada como $\Delta(z) = z \delta / g_r$, donde δ es el tamaño del píxel. Por tanto, podemos considerar que el espacio 3D ante el monitor de InI está dividido en sectores, cada uno de los cuales se corresponde a la proyección de un píxel del monitor de InI a través

de la lente central, ver Figura 7.2. Bajo esta consideración, cuando el observador se mueve de un sector a otro, el software de control modifica las microimágenes que se muestran en el monitor de InI. En concreto, si el observador se encuentra justo enfrente del monitor de InI, en el sector central, las microimágenes no se encontrarán desplazadas (como en un monitor de InI convencional), ver Figura 7.1 (a). Sin embargo, si el observador se encuentra en cualquier otro sector, la máscara de recorte y las microimágenes que se muestran en el monitor serán desplazadas un número de píxeles igual al número de sectores entre el sector central y en el que se encuentre el observador (inclusive), ver Figura 7.1 (b). Nótese que las microimágenes siempre son desplazadas en dirección contraria a la que el observador se ha movido.

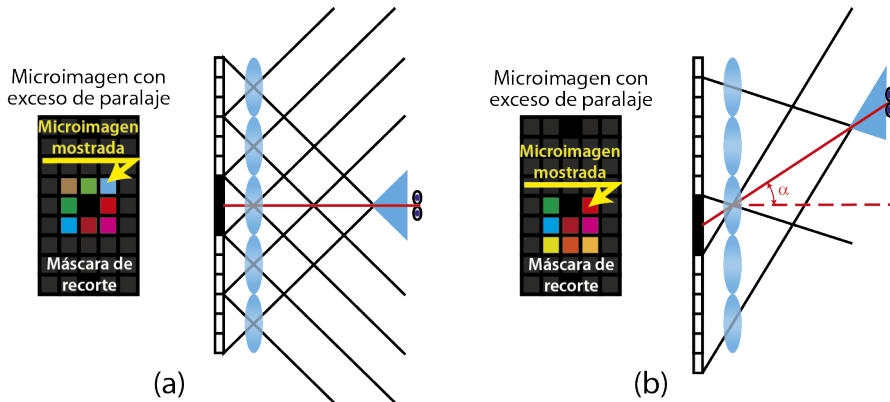


Figura 7.1.: (a) Representación esquemática del proceso de recorte y desplazamiento de las microimágenes y el AV cuando el observador se coloca justo enfrente de la microimagen central; (b) Representación esquemática del proceso de recorte y desplazamiento de las microimágenes y AV cuando el observador se desplaza con respecto al centro del monitor de InI.

Cambio dinámico de la posición del PR

Finalmente, para mejorar aún más la experiencia de visualización, nuestro método también permite al observador cambiar el PR a voluntad. Para ello, el software de control modifica el PR de la imagen plenóptica

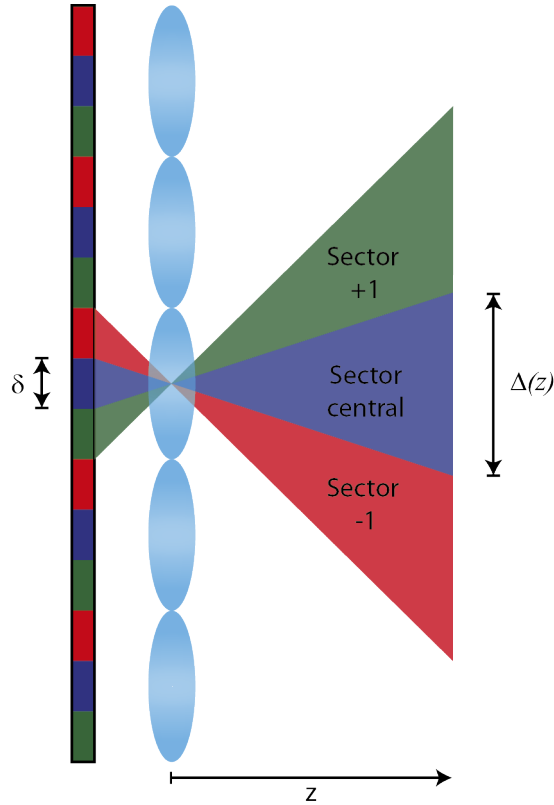


Figura 7.2.: Representación de los sectores en los que se divide el espacio enfrente del monitor de InI. Cada sector se corresponde con la proyección de un pixel a través de la lente central.

mostrada en el monitor de InI dependiendo de la distancia entre el observador y el monitor. La ventaja de este añadido es que el observador puede controlar con su posicionamiento qué partes de la escena 3D serán reconstruidas delante o detrás de la MR. Para cambiar el PR de la imagen plenótica, usamos el algoritmo de SPOC 2.0 [65].

7.2 Resultados experimentales

Para ilustrar nuestra propuesta, hemos aplicado nuestro método a un monitor de InI. Para no complicar en exceso el experimento, decidimos

incrementar únicamente el paralaje horizontal de la imagen 3D producida.

En primer lugar implementamos el monitor de InI. Hasta el momento, en esta Tesis hemos utilizado un iPad y una MR pequeña. En esta ocasión, hemos utilizado un monitor 4K de 24 pulgadas (DellP2415Q) con 3840×2160 píxeles (resolución 7,29 píxeles/mm) como pantalla y una MR compuesta por 327×184 lentes de tamaño $p_r = 1,607$ mm y distancia focal $f_r = 8$ mm. Como vimos en la sección 4.2, podemos reducir el efecto del trenzado de facetas, sin enfatizar en exceso el efecto subpixel, si situamos los píxeles en un plano ligeramente desplazado respecto al plano focal de las lentes. Por tanto, fijamos la separación entre los píxeles y las lentes en $g_r = 8,4$ mm (véase anexo A). Con ello, disminuimos el efecto subpixel sin sacrificar calidad en la imagen 3D producida.

Para realizar la captura de la imagen integral, volvemos a utilizar el método de la apertura sintética [51]. El número de píxeles por microimagen en nuestro monitor de InI, 12×12 , es igual al número de perspectivas que éste puede mostrar simultáneamente. Entonces, para tener un exceso de vistas y paralaje, capturamos un total de 40×12 EIs. La configuración de captura que utilizamos es similar a la empleada en los experimentos realizados en capítulos anteriores. Usamos una cámara Canon 450D montada sobre unos motores cuya posición y desplazamiento son controlados con Labview[®]. Establecimos una separación lateral entre fotos de $P_H = P_V = 5$ mm y utilizamos una escena 3D compuesta por una casa de Lego[®], un muñeco y una caja situados a 39, 48 y 55 mm de la cámara, respectivamente. Para los parámetros de la cámara usamos una focal de $f = 18$ mm y un número- f de $f/22$ (lo suficientemente grande para conseguir que la escena 3D sea capturada nítidamente).

El siguiente paso consiste en calcular, a partir de las EIs capturadas, las microimágenes con exceso de paralaje. Para ello, usamos el algoritmo SPOC 2.0 [65], con el que obtenemos una imagen plenóptica compuesta

por 230×184 microimágenes de 40×12 píxeles cada una. A partir de esta imagen plenóptica, y aplicando el correspondiente recorte para cada posible posición lateral del observador, calculamos 28 imágenes plenópticas compuestas por microimágenes de 12×12 píxeles. Finalmente, para adaptar las imágenes al monitor de InI, ajustamos su tamaño para que las microimágenes tengan $11,7 \times 11,7$ píxeles cada una. A continuación, repetimos este proceso y calculamos otro conjunto de 28 imágenes plenópticas pero, en este caso, escogemos una segunda posición para el PR de la imagen plenóptica.

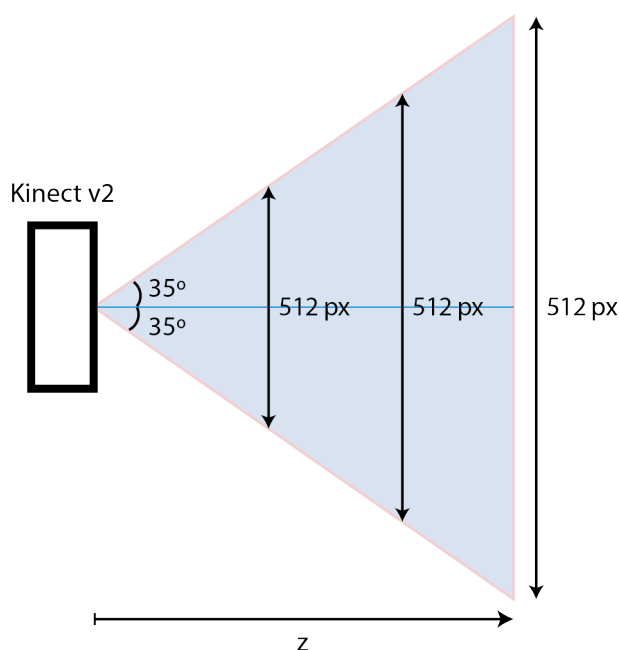


Figura 7.3.: Área de seguimiento del Kinect v2. El número de píxeles (px en la imagen) del sensor IR es 512.

Para el sistema eye-tracking usamos un dispositivo Kinect v2, centrado lateralmente con el monitor de InI, en conjunto con un kit de desarrollo de software de Microsoft (del inglés *Microsoft SDK*) que permite al Kinect seguir la posición de observadores en movimiento. El Kinect v2 funciona con tecnología de tiempo de vuelo (del inglés *time-of-flight*) IR, por lo que puede obtener mapas de profundidad de elevada precisión y operar en

condiciones de baja luminosidad [93]-[95]. Para realizar el seguimiento de un observador, este dispositivo localiza la posición de la cabeza y las pupilas del observador. No obstante, para que funcione, el observador debe moverse en el interior del campo visual y el rango de seguimiento en profundidad del Kinect. Los ángulos de visualización del Kinect son $70^\circ(H) \times 60^\circ(V)$. Aunque la región de seguimiento en profundidad de un dispositivo Kinect v2 cubre desde 500 mm hasta 8000 mm, el intervalo en el cual se puede obtener la posición del observador con precisión óptima se reduce a 1000-4500 mm aproximadamente. En este rango óptimo, la precisión para la determinación de la posición lateral de un determinado objeto (como la pupila del ojo) es de $0,137^\circ$. Por otro lado, la precisión en la distancia en profundidad es de 2 mm [95]. Algunas de las características del Kinect v2 se ilustran en la Figura 7.3.

A continuación, hallamos la correspondencia entre los 512 píxeles del sensor IR del Kinect v2 y los sectores en los que dividimos el espacio enfrente del monitor de InI. Para ello, necesitamos conocer la proyección de un píxel del sensor en mm. Esta proyección depende de la distancia y puede ser calculada como:

$$AS = 2z \tan(35^\circ), \quad (7.1)$$

$$IR_{mm} = \frac{AS}{IR_{pix}}, \quad (7.2)$$

donde AS (área de seguimiento) es el campo visual del Kinect v2 en mm, IR_{pix} se corresponde con el número de píxeles del sensor IR, IR_{mm} representa el tamaño de la proyección de un píxel del sensor IR en la posición del observador, y z es la separación entre el monitor de InI y el observador. Entonces, el tamaño, S , de cada sector expresado en unidades de píxeles del sensor IR, es:

$$S = \frac{\Delta(z)}{IR_{mm}}. \quad (7.3)$$

Usar las ecuaciones 7.1-7.3 proporciona un resultado interesante: el tamaño de cada sector en unidades de píxeles IR es 6 píxeles y es independiente de la distancia del observador al monitor de InI. Por tanto, podemos dividir el sensor IR del Kinect v2 en conjuntos de 6 píxeles y asignarle a cada conjunto su correspondiente imagen regenerada.

Para llevar a cabo el experimento, el software de control selecciona la imagen plenóptica que se mostrará en el monitor de InI dependiendo de la posición 3D del observador. Cuando el Kinect v2 le comunica al software que el observador se ha movido de un conjunto de 6 píxeles a otro conjunto de 6 píxeles diferente, el software de control cambia la imagen plenóptica mostrada por la siguiente. Asimismo, el software también cambiará la imagen mostrada por otra con un PR diferente en función de la posición en profundidad del observador. Como hemos preparado imágenes con dos PR diferentes, hemos seleccionado dos planos arbitrarios a distinta distancia del monitor de InI. El primer plano se encuentra a una distancia de $z = 1940$ mm y el segundo plano a $z = 2040$ mm. El PR de la imagen que se muestra en el monitor dependerá de en cual de estos dos planos se sitúe el observador. Nótese que el motivo por el que elegimos un rango de 100 mm para realizar el cambio es porque un desplazamiento en profundidad de esa cantidad es suficiente para proporcionar una buena experiencia de visualización al observador.

Para demostrar la utilidad de nuestro método hemos grabado 5 vídeos. En **VÍDEO 7.1** y **VÍDEO 7.2** comparamos nuestra propuesta con un monitor de InI operando de forma convencional. En ambos casos, el observador se mueve lateralmente enfrente del monitor de InI a una distancia de $z = 2040$ mm. Para este valor de z , el tamaño del área de visualización es de $AV = 16$ mm. Por tanto, la restricción sobre la posición lateral del observador es bastante rígida si el monitor de InI está trabajando en modo

convencional. En este caso, el ángulo de visualización sería de $\alpha = 0,3^\circ$. En **VÍDEO 7.1**, usamos una configuración de monitor de InI convencional, sin seguimiento y mostrando siempre la misma imagen. Por otro lado, **VÍDEO 7.2** se corresponde con la técnica propuesta donde el eye-tracking está activado y la imagen mostrada en el monitor se adapta a la posición del observador. Teniendo en cuenta el tamaño de las microimágenes sin recortar (40×12), el ángulo de visualización aumenta, en este caso, a $\alpha = 14,8^\circ$. Entonces, el observador puede cambiar su posición lateral a lo largo de un intervalo superior a 1000 mm.

Para grabar los vídeos, usamos una cámara digital montada en una plataforma móvil para simular los movimientos de un observador. En ambos experimentos el desplazamiento lateral fue de 800 mm aproximadamente. A partir de los vídeos es aparente que nuestro método permite al observador realizar grandes desplazamientos en paralelo a la pantalla sin percibir crosstalk. Asimismo, el paralaje se ha incrementado notablemente (nótese que en **VÍDEO 7.1** el observador percibe en repetidas ocasiones las mismas perspectivas). En la Figura 7.4 enseñamos cuatro fotogramas de **VÍDEO 7.1** y **VÍDEO 7.2**.

VÍDEO 7.3 muestra los resultados experimentales para un observador moviéndose hacia el monitor de InI y alejándose del mismo. Para grabar este vídeo también usamos una cámara para simular los movimientos del observador. Como se aprecia en el vídeo, cuando el observador se encuentra lejos del monitor, el PR se sitúa al frente de la caja (azul y amarilla), mientras que el resto de la escena se reconstruye detrás de la MR. Se puede ver que la mayor calidad se obtiene en el PR y, por tanto, otras partes de la escena 3D (como la casa) aparecen ligeramente desenfocadas. Más adelante en el vídeo, cuando el observador se acerca al monitor, el PR se sitúa en la fachada de la casa. En este caso, el resto de la escena 3D se reconstruye por delante de la MR, y la caja es percibida con algo de desenfoque. El efecto de aumento que se puede percibir en el

vídeo se produce debido al hecho de que cuando nos aproximamos a la escena ésta es percibida con mayor tamaño angular.

En **VÍDEO 7.4** enseñamos como el monitor de InI adapta su imagen, en tiempo real, al desplazamiento lateral de un observador humano. Nótese que el vídeo no está grabado desde el punto de vista del observador, sino desde otra perspectiva. Por tanto, lo que se ve en el vídeo no es la imagen como sería percibida por el observador (que se puede ver en **VÍDEO 7.2**). Análogamente, en **VÍDEO 7.5** enseñamos como el monitor de InI adapta la imagen mostrada a los cambios en la posición de profundidad del observador humano. De nuevo, este vídeo muestra una perspectiva diferente a la percibida por el observador (que se muestra en **VÍDEO 7.3**). Entonces, el efecto crosstalk que aparece en ambos vídeos es sólo percibido por la cámara utilizada para grabar, no por el observador. En la Figura 7.5 enseñamos un fotograma extraído de **VÍDEO 7.5**.



Figura 7.4.: Comparación entre nuestro método (columna izquierda) con una configuración convencional del monitor de InI (columna derecha). Las imágenes fueron capturadas en diferentes posiciones laterales enfrente del monitor. El desplazamiento total de la cámara fue aproximadamente de 800 mm. En la configuración convencional, el crosstalk puede ser visto en el lado izquierdo de la casa (segunda imagen) y sobre el lado derecho de la casa (tercera imagen).

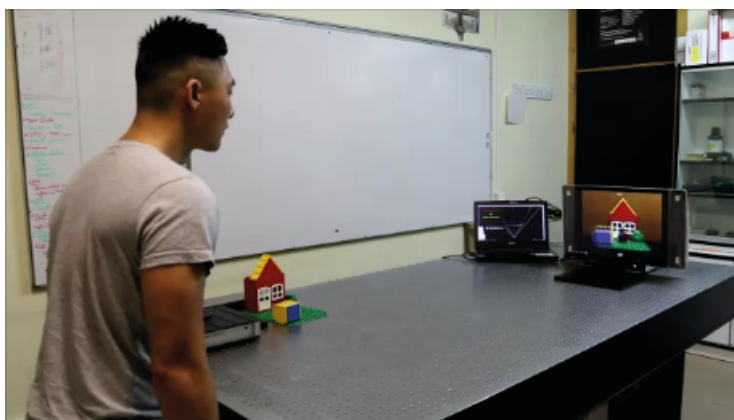


Figura 7.5.: Fotografía del montaje experimental y del proceso de grabación de **VÍDEO 7.5**. El dispositivo utilizado está compuesto por un monitor de InI, un Kinect v2, y un ordenador que recibe la información del Kinect y transmite al monitor la correspondiente imagen plenóptica. La imagen que se muestra en el monitor de InI se adapta a la posición 3D del observador en tiempo real.

Conclusiones

En esta Tesis se han estudiado las características de los sistemas de InI. Con ello, se han podido analizar los mecanismos mediante los cuales un monitor de InI produce imágenes 3D y plantear las características que uno de estos dispositivos debería tener para producir imágenes de calidad: un monitor de alta resolución y una matriz con lentes lo más pequeñas posibles.

Asimismo, se han analizado las técnicas de captura de InI y de fotografía plenóptica. Aunque ambas técnicas difieren en su implementación, se ha visto que existe una relación de equivalencia entre ambas. A partir de esta equivalencia, se ha diseñado un método que permite convertir las EIs, obtenidas con un sistema de InI, en microimágenes, que son más adecuadas para ser mostradas en el monitor de InI planteado. Asimismo, las imágenes 3D producidas con las microimágenes obtenidas usando esta técnica presentan mayor paralaje y pueden formar imágenes 3D de escenas más grandes que las obtenidas mediante otros métodos más convencionales. Por otro lado, este método también permite adaptar las microimágenes generadas al monitor de InI en el que se van a mostrar y seleccionar a voluntad varias de las características que tendrá la imagen 3D generada, pudiendo generar diferentes experiencias de visualización a partir de la misma imagen integral.

Aprovechando esta última particularidad del método propuesto, se ha diseñado una aplicación para el mismo, consistente en introducir una rutina en el algoritmo utilizado que permite producir un conjunto de imágenes plenópticas que, al ser reproducidas en secuencia, conforman

un vídeo que muestra un trávelin de la imagen 3D. Por otra parte, se ha propuesto una segunda aplicación para dicho método, cuyo objetivo es solventar uno de los principales inconvenientes de un monitor de InI; su escaso ángulo de visualización. Para ello, se alteró el algoritmo del método para que fuera capaz de producir microimágenes de una distribución específica. De este modo, se pudo obtener un conjunto de microimágenes con distribución y tamaño ajustado a una MR girada, lo que permite aumentar el ángulo de visualización de la imagen 3D producida.

Finalmente, teniendo en cuenta todas las propuestas anteriores, se ha diseñado un sistema que permite mejorar el ángulo de visualización y el paralaje que puede mostrar un monitor de InI. Este sistema adapta la información mostrada en el monitor de InI, en tiempo real, a la posición 3D del observador y, además, permite al observador elegir qué partes de la imagen 3D se forman delante y detrás de la MR. El monitor de InI implementado en esta última propuesta satisface los requisitos del objetivo principal de esta Tesis: implementar un monitor de InI con componentes comerciales que produzca imágenes 3D de calidad de escenas reales, que proporcione una buena experiencia de visualización al observador y que sirva como prototipo para un sistema de captura y display de imágenes 3D en tiempo real. No obstante, hay que destacar que aún se necesita bastante trabajo para conseguir que este prototipo evolucione a un sistema de InI que funcione en tiempo real, siendo la falta de optimización en la captura de la imagen integral, la transmisión de información y el tiempo de ejecución del algoritmo del método propuesto unas de las principales trabas.

Como vimos en el capítulo 4, la separación, g_r , entre los píxeles de la pantalla y las lentes de la MR de un monitor afecta a las características de la imagen 3D producida. A continuación, analizamos esta distancia en los dos monitores de InI implementados en esta Tesis.

Matriz de lentes plano-convexa.

Las lentes de las matrices empleadas en ambos monitores de InI son plano-convexas y tienen un grosor no despreciable (3,3 mm y 4 mm aproximadamente). En consecuencia, la aproximación de lente delgada no es adecuada. Por tanto, para determinar correctamente la distancia g_r , necesitamos representar las lentes de la matriz a partir de sus planos principales. Para ello, calculamos los valores de h_1 y h_2 que se corresponden con las distancias $\overline{V_1H_1}$ y $\overline{V_2H_2}$, respectivamente (ver Figura A.1) [40]:

$$h_1 = -\frac{f(n_l - 1)d_l}{R_2n_l} \quad (\text{A.1})$$

$$h_2 = -\frac{f(n_l - 1)d_l}{R_1n_l} \quad (\text{A.2})$$

En estas ecuaciones, d_l , n_l y f representan, respectivamente, la anchura, el índice de refracción y la focal de la lente. R_1 y R_2 son los radios de curvatura de las superficies de la lente. En nuestro caso, cuando

implementamos el monitor de InI, orientamos la MR para que la luz incida sobre la cara plana de las lentes. Por tanto, teniendo en cuenta que $R_1 = 0$ y $1/f = -(n_l - 1)/R_2$ [40], obtenemos unos valores de $h_2 = 0$ y $h_1 = d_l/n_l$.

De manera alternativa, también podemos obtener las posiciones de los planos principales de forma experimental. Para ello, basta con hacer incidir un haz de luz paralelo sobre la lente y anotar la posición dónde la luz se focaliza. Al hacer esto por ambas superficies de la lente se obtendrá la posición de los dos planos focales. Después, podemos calcular la posición de los planos principales teniendo en cuenta que $\overline{F_1 H_1} = f$ y $\overline{F_2 H_2} = -f$.

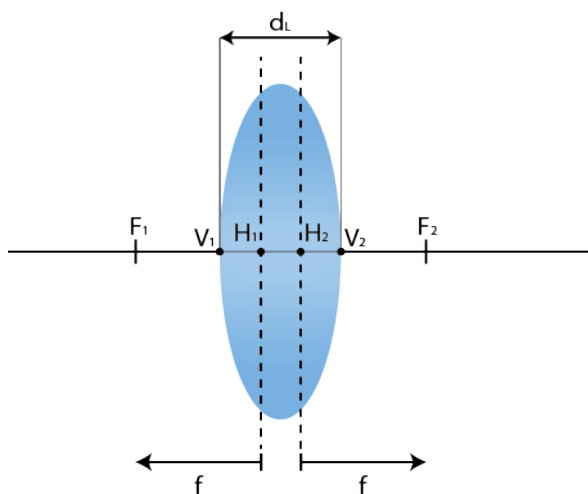


Figura A.1.: Representación esquemática de una lente gruesa. Las dos líneas discontinuas se corresponden con los planos principales de la lente.

Teniendo en cuenta las características de las matrices de lentes utilizadas, si quisiéramos situar los píxeles del monitor en el plano focal de las lentes de la MR, éstos deberían encontrarse a una distancia aproximada (de la cara plana de las lentes) de $g_r = 1,09$ mm para la matriz de lentes utilizada en el iPad ($f_r = 3,3$ mm y $p_r = 1$ mm) y $g_r = 5,0$ mm para la utilizada en el monitor 4K ($p_r = 1,607$ y $f_r = 8$ mm).

Película protectora de los píxeles.

Hay que tener en cuenta que es habitual que exista una película o capa protectora entre los píxeles de una pantalla y su superficie. Por tanto, la anchura, k , de ésta también tendrá que ser tenida en cuenta para determinar correctamente la distancia g_r . Para medir k , en esta Tesis hemos utilizado un microscopio que nos permitió enfocar fácilmente a los planos dónde se encontraban los píxeles y la superficie de la capa protectora. En el caso del iPad, la anchura medida es de, aproximadamente, $k = 1,5$ mm, mientras que para el monitor 4K es $k = 0,4$ mm.

Listado de vídeos

Enlace

- 5.1 <https://nuvol.uv.es/owncloud/index.php/s/j9vS6nDoEDOp3n8>
- 5.2 <https://nuvol.uv.es/owncloud/index.php/s/zZckGC7ZaertGMs>
- 5.3 <https://nuvol.uv.es/owncloud/index.php/s/LZIj9dsXFNQYlaF>
- 5.4 <https://nuvol.uv.es/owncloud/index.php/s/FzVUlxxqnbJaWAD>
- 6.1 <https://nuvol.uv.es/owncloud/index.php/s/pE5tLMcCLWgjMb2>
- 6.2 <https://nuvol.uv.es/owncloud/index.php/s/PYGjzTYHhB6TVsg>
- 6.3 <https://nuvol.uv.es/owncloud/index.php/s/NDKou2gaYuuLu42>
- 6.4 <https://nuvol.uv.es/owncloud/index.php/s/CwleoPzFAd8DY5E>
- 6.5 <https://nuvol.uv.es/owncloud/index.php/s/Ybl0wjDXUDZb86u>
- 7.1 <https://nuvol.uv.es/owncloud/index.php/s/Ncwlp9AYaye0Q95>
- 7.2 <https://nuvol.uv.es/owncloud/index.php/s/1Xzx42aIprakjFx>
- 7.3 <https://nuvol.uv.es/owncloud/index.php/s/A6Em2sJk2EN4jMs>

7.4 <https://nuvol.uv.es/owncloud/index.php/s/5YpHUEVpwfDOAY1>

7.5 <https://nuvol.uv.es/owncloud/index.php/s/z2MmWTY1IRasaBW>

Bibliografía

- [1]C. Wheatstone, “XVIII. Contributions to the physiology of vision. —Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision”, *Philos. T. R. Soc. Lond*, vol. 128, págs. 371-394, 1838 (vid. pág. 1).
- [2]W. Rollmann, “Notiz zur Stereoskopie”, *Ann. Phys.*, vol. 165, n.º 6, págs. 350-351, 1853 (vid. pág. 1).
- [3]E. H. Land, *Polarizing optical system*, *U. S. Patent 2.099.694*, 1937 (vid. pág. 1).
- [4]S. S. Kim, B. H. You, H. Choi y col., “World’s First 240Hz TFT-LCD Technology for Full-HD LCD-TV and Its Application to 3D Display”, *SID Symposium Digest of Technical Papers*, vol. 40, n.º 1, pág. 424, 2009 (vid. pág. 1).
- [5]W. Hess, “Stereoscopic picture”, *U. S. Patent 1,128,979*. 1915 (vid. pág. 1).
- [6]F. E. Ives, “A novel stereogram”, *J. Frankl. Inst.*, vol. 153, págs. 51-52, 1902 (vid. pág. 1).
- [7]T. Inoue y H. Ohzu, “Accommodative responses to stereoscopic three-dimensional display”, *Appl. Opt.*, vol. 36, n.º 19, págs. 4509-4515, 1997 (vid. págs. 1, 52).
- [8]S. Yano, S. Ide, T. Mitsuhashi y H. Thwaites, “A study of visual fatigue and visual comfort for 3D HDTV/HDTV images”, *Displays*, vol. 23, n.º 4, págs. 191-201, 2002 (vid. págs. 1, 52).
- [9]F. L. Kooi y A. Toet, “Visual comfort of binocular and 3D displays”, *Displays*, vol. 25, n.º 2-3, págs. 99-108, 2004 (vid. págs. 1, 52).

- [10]H. Hiura, K. Komine, J. Arai y T. Mishina, "Measurement of static convergence and accommodation responses to images of integral photography and binocular stereoscopy.", *Opt. express*, vol. 25, n.º 4, págs. 3454-3468, 2017 (vid. págs. 1, 52).
- [11]G. Lippmann, "Épreuves Réversibles Donnant la Sensation du Relief", *J. Phys.*, vol. 7, n.º 1, págs. 821-825, 1908 (vid. págs. 1, 2, 6, 18, 39).
- [12]J.-H. Park, K. Hong y B. Lee, "Recent progress in three-dimensional information processing based on integral imaging", *Appl. Opt.*, vol. 48, n.º 34, H77-H94, 2009 (vid. págs. 1, 3, 75).
- [13]T. Okoshi, "Three-Dimensional Displays", *Proc. IEEE*, vol. 68, n.º 5, págs. 548-564, 1980 (vid. págs. 1).
- [14]C. Slinger, C. Cameron y M. Stanley, "Computer-generated holography as a generic display technology", *Computer*, vol. 38, n.º 8, págs. 46-53, 2005 (vid. págs. 1).
- [15]B. Ridwan, T. Adrian, X. Xuewu y col., "Digital holographic three-dimensional display of 50-Mpixel holograms using a two-axis scanning mirror device", *J. Opt. Eng.*, vol. 49, n.º 2, págs. 025 801, 2010 (vid. págs. 1).
- [16]H. E. Ives, "Motion Pictures in Relief", *J. Opt. Soc. Am.*, vol. 18, n.º 2, págs. 118-122, 1929 (vid. págs. 1).
- [17]—, "The Projection of Parallax Panoramagrams", *J. Opt. Soc. Am.*, vol. 21, n.º 7, págs. 397-409, 1931 (vid. págs. 1).
- [18]—, "Optical Properties of a Lippmann Lenticulated Sheet", *J. Opt. Soc. Am.*, vol. 21, n.º 3, págs. 171-176, 1931 (vid. págs. 1, 74, 75).
- [19]—, "Parallax Panoramagrams Made with a Large Diameter Lens", *J. Opt. Soc. Am.*, vol. 20, n.º 6, págs. 332-342, 1930 (vid. págs. 2).
- [20]D. F. Coffey, "Apparatus for Making a Composite Stereograph", *U.S. Patent 2,063,985* (1936) (vid. págs. 2).
- [21]E. H. Adelson y J. Y. A. Wang, "Single Lens Stereo with a Plenoptic Camera", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, n.º 2, págs. 99-106, 1992. arXiv: 0070242542 (vid. págs. 2, 7, 25).
- [22]F. Okano, H. Hoshino, J. Arai e I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography.", *Appl. Opt.*, vol. 36, n.º 7, págs. 1598-1603, 1997 (vid. págs. 2, 75).

- [23]R. Ng, M. Levoy, G. Duval, M. Horowitz y P. Hanrahan, “Light Field Photography with a Hand-held Plenoptic Camera”, *Informational*, págs. 1-11, 2005. arXiv: 0070242542 (vid. págs. 3, 8, 24, 30).
- [24]H. Navarro, “Captura y reproducción de imágenes 3D mediante sistemas de Imagen Integral. Diseño, implementación y aplicaciones.”, Tesis doct., Universidad de Valencia, 2014 (vid. págs. 3, 8, 34, 52, 54, 75).
- [25]A. Llavador, E. Sánchez-Ortiga, G. Saavedra, B. Javidi y M. Martínez-Corral, “Free-depths reconstruction with synthetic impulse response in integral imaging”, *Opt. Express*, vol. 23, n.º 23, págs. 30 127-30 135, 2015 (vid. pág. 3).
- [26]A. Llavador, “Captura y procesado de imágenes de alta resolución mediante sistemas de microscopía integral”, Tesis doct., Universidad de Valencia, 2017 (vid. págs. 3, 8, 34).
- [27]S. Hong y B. Javidi, “Three-dimensional visualization of partially occluded objects using integral imaging”, *J. Disp. Technol*, vol. 1, n.º 2, págs. 354-359, 2005 (vid. pág. 3).
- [28]H. Navarro, G. Saavedra, A. Molina y col., “Optical slicing of large scenes by synthetic aperture integral imaging”, *Proc. SPIE*, págs. 7690-0M, 2010 (vid. pág. 3).
- [29]O. Matoba, E. Tajahuerce y B. Javidi, “Real-time three-dimensional object recognition with multiple perspectives imaging.”, *Appl. Opt.*, vol. 40, n.º 20, págs. 3318-3325, 2001 (vid. pág. 3).
- [30]M. Levoy, R. Ng, A. Adams, M. Footer y M. Horowitz, “Light field microscopy”, *ACM Transactions on Graphics*, vol. 25, n.º 3, págs. 924-934, 2006. arXiv: 1508.03590 (vid. págs. 3, 30).
- [31]A. Llavador, J. Sola-Pikabea, G. Saavedra, B. Javidi y M. Martínez-Corral, “Resolution improvements in integral microscopy with Fourier plane recording”, *Opt. Express*, vol. 24, n.º 18, pág. 20 792, 2016 (vid. pág. 3).
- [32]H. Navarro, M. Martínez-Corral, G. Saavedra, A. Pons y B. Javidi, “Photoelastic analysis of partially occluded objects with an integral-imaging polariscope”, *J. Disp. Technol*, vol. 10, n.º 4, págs. 255-262, 2014 (vid. pág. 3).
- [33]S. Hong, D. Shin, B. G. Lee y col., “Towards 3D Television Through Fusion of Kinect and Integral-Imaging Concepts”, *J. Disp. Technol*, vol. 11, n.º 11, págs. 894-899, 2015 (vid. pág. 3).

- [34]X. Xiao, B. Javidi, M. Martínez-Corral y A. Stern, “Advances in three-dimensional integral imaging: sensing, display, and applications”, *Appl. Opt.*, vol. 52, n.º 4, págs. 546-560, 2013 (vid. pág. 3).
- [35]M. Martínez-Corral, A. Dorado, J. C. Barreiro, G. Saavedra y B. Javidi, “Recent Advances in the Capture and Display of Macroscopic and Microscopic 3-D Scenes by Integral Imaging”, *Proc. IEEE*, vol. 105, n.º 5, págs. 825-836, 2017 (vid. pág. 3).
- [36]F. Okano, J. Arai, K. Mitani y M. Okui, “Real-time integral imaging based on extremely high resolution video system”, *Proc. IEEE*, vol. 94, n.º 3, págs. 122-125, 2006 (vid. pág. 3).
- [37]J. Arai, F. Okano, M. Kawakita y col., “Integral three-dimensional television using a 33-megapixel imaging system”, *J. Disp. Technol*, vol. 6, n.º 10, págs. 422-430, 2010 (vid. pág. 3).
- [38]J. Arai, “Three-dimensional television system based on spatial imaging method using integral photography”, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, n.º c, págs. 5449-5452, 2012 (vid. pág. 3).
- [39]J. Kim, J.-H. Jung, C. Jang y B. Lee, “Real-time capturing and 3D visualization method based on integral imaging”, *Opt. Express*, vol. 21, n.º 16, pág. 18 742, 2013 (vid. págs. 3, 55, 73).
- [40]E. Hecht, *Óptica*, 5º. Pearson Educación, S. A., 2017 (vid. págs. 5, 125, 126).
- [41]M. Martínez-Corral, W. Furlan, A. Pons y G. Saavedra, *Instrumentos Ópticos y Optométricos. Teoría y prácticas*. Universitat de Valencia. Servei de publicacions, 1998 (vid. págs. 7, 16).
- [42]E. H. Adelson y J. R. Bergen, “The plenoptic function and the elements of early vision”, *Computational models of visual processing*, vol. 1, págs. 3-20, 1991 (vid. pág. 7).
- [43]A. Gershun, “The Light Field”, *Journal Of Mathematics And Physics*, vol. 18, n.º 2, págs. 51-151, 1939 (vid. pág. 7).
- [44]M. Levoy y P. Hanrahan, “Light Field Rendering”, *Proc. SIGGRAPH, ACM*, págs. 31-42, 1996 (vid. págs. 7, 8).
- [45]M. Levoy, “Light Fields and Computational Imaging”, *Computer*, vol. 39, n.º 8, págs. 46-55, 2006 (vid. pág. 7).

- [46]A. Lumsdaine y T. Georgiev, "The focused plenoptic camera", *Proc. IEEE International Conference on Computational Photography*, págs. 1-8, 2009 (vid. págs. 8, 24, 31).
- [47]R. Ng, "Digital light field photography", Tesis doct., 2006, págs. 1-203 (vid. págs. 8, 24).
- [48]K. Halbach, "Matrix Representation Of Gaussian Optics", *American Journal of Physics*, vol. 32, n.º 2, págs. 90-108, 1964 (vid. págs. 12).
- [49]J. Casas Peláez, *Optica* (vid. págs. 12).
- [50]M. Martínez-Corral, R. Martínez-Cuenca, G. Saavedra y col., "Progresses in 3D integral imaging with optical processing", *Journal of Physics: Conference Series*, vol. 139, 2008 (vid. págs. 22, 46).
- [51]J.-S. Jang y B. Javidi, "Three-dimensional synthetic aperture integral imaging", *Opt. Letters*, vol. 27, n.º 13, págs. 1144, 2002 (vid. págs. 23, 81, 96, 102, 115).
- [52]Y. S. Hwang, S. H. Hong y B. Javidi, "Free view 3-D visualization of occluded objects by using computational synthetic aperture integral imaging", *J. Disp. Technol*, vol. 3, n.º 1, págs. 64-70, 2007 (vid. págs. 23).
- [53]H. Navarro, M. Martínez-Corral, A. Dorado y col., "Capture of the spatio-angular information of a 3D scene", *Optica Pura y Aplicada*, vol. 46, n.º 2, 2013 (vid. págs. 26).
- [54]H. Navarro, J. C. Barreiro, G. Saavedra, M. Martínez-Corral y B. Javidi, "High-resolution far-field integral-imaging camera by double snapshot", *Opt. Express*, vol. 20, n.º 2, págs. 890-895, 2012 (vid. págs. 26, 30, 35, 77).
- [55]A. Lumsdaine, "Focused plenoptic camera and rendering", *J. Electron. Imaging*, vol. 19, n.º 2, págs. 021 106, 2010 (vid. págs. 31).
- [56]R. Martínez-Cuenca, "Captura y reproducción de imágenes tridimensionales mediante sistemas multiperspectiva", Tesis doct., Universidad de Valencia (vid. págs. 47, 69).
- [57]Y. Kim, J. Kim, K. Hong y col., "Accommodative response of integral imaging in near distance", *J. Disp. Technol*, vol. 8, n.º 2, págs. 70-78, 2012 (vid. págs. 52).
- [58]S. Y. H. Hiura, T. Mishina, J. Arai, K. Hisatomi, Y. Iwadate, T. Ito, "A study on accommodation response and depth perception in viewing integral photography", *Proc. 3DSA*, vol. 2, págs. 2, 2013 (vid. págs. 52).

- [59]R. Martínez-Cuenca, G. Saavedra, A. Pons, B. Javidi y M. Martínez-Corral, "Facet braiding: a fundamental problem in integral imaging.", *Opt. Lett.*, vol. 32, n.º 9, págs. 1078-1080, 2007 (vid. págs. 52, 73).
- [60]H. Navarro, R. Martínez-Cuenca, A. Molina-Martían y col., "Method to remedy image degradations due to facet braiding in 3D integral-imaging monitors", *J. Disp. Technol*, vol. 6, n.º 10, págs. 404-411, 2010 (vid. págs. 52, 54, 55).
- [61]M. Okui, M. Kobayashi, J. Arai y F. Okano, "Moire fringe reduction by optical filters in integral three-dimensional imaging on a color flat-panel display", *Appl. Opt.*, vol. 44, n.º 21, págs. 4475-4483, 2006 (vid. págs. 55).
- [62]Y. Kim, G. Park, S.-W. Cho y col., "Integral imaging with reduced color moire pattern by using a slanted lens array", *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 6803, págs. 68031L-6803-8, 2008 (vid. págs. 55).
- [63]Y. Kim, G. Park, J.-H. Jung, J. Kim y B. Lee, "Color moiré pattern simulation and analysis in three-dimensional integral imaging for finding the moiré-reduced tilted angle of a lens array", *Appl. Opt.*, vol. 48, n.º 11, págs. 2178, 2009 (vid. págs. 55).
- [64]Y. Kim, J. Yeom, J.-h. Jung, J. Hong y B. Lee, "View image error analysis based on focal mode and virtual mode in three-dimensional display using lenses", *Advances*, vol. 7956, n.º February 2011, págs. 1-6, 2011 (vid. págs. 55).
- [65]M. Martínez-Corral, A. Dorado, H. Navarro, G. Saavedra y B. Javidi, "Three-dimensional display by smart pseudoscopic-to-orthoscopic conversion with tunable focus", *Appl. Opt.*, vol. 53, n.º 22, E19-E26, 2014 (vid. págs. 65, 80, 89, 93, 109, 114, 115).
- [66]*Lightfield based commercial digital still camera*. <http://www.lytro.com> (vid. págs. 69).
- [67]*3D lightfield camera*. <http://www.raytrix.de/> (vid. págs. 69).
- [68]A. Tolosa, A. Pons, G. Saavedra y B. Javidi, "Optical implementation of micro-zoom arrays for parallel focusing in integral imaging", *America*, vol. 27, n.º 3, págs. 495-500, 2010 (vid. págs. 73).
- [69]J.-H. Jung, J. Kim y B. Lee, "Solution of pseudoscopic problem in integral imaging for real-time processing.", *Opt. Lett.*, vol. 38, n.º 1, págs. 76-8, 2013 (vid. págs. 73, 75).

- [70]M. Martínez-Corral, B. Javidi, R. Martínez-Cuenca y G. Saavedra, "Formation of real, orthoscopic integral images by smart pixel mapping", *Opt. Express*, vol. 13, n.º 23, págs. 9175-9180, 2005 (vid. pág. 75).
- [71]A. Aggoun, "Pre-processing of integral images for 3-D displays", *J. Disp. Technol*, vol. 2, n.º 4, págs. 393-400, 2006 (vid. pág. 75).
- [72]K. S. Park, S. W. Min e Y. Cho, "Viewpoint vector rendering for efficient elemental image generation", *IEICE Transactions on Information and Systems*, vol. E90-D, n.º 1, págs. 233-241, 2007 (vid. pág. 75).
- [73]D. H. Shin, B. G. Lee y E. S. Kim, "Modified smart pixel mapping method for displaying orthoscopic 3D images in integral imaging", *Optics and Lasers in Engineering*, vol. 47, n.º 11, págs. 1189-1194, 2009 (vid. pág. 75).
- [74]E. Sahin y L. Onural, "A comparative study of light field representation and integral imaging", *Imaging Science Journal*, vol. 58, n.º 1, págs. 28-31, 2010 (vid. pág. 75).
- [75]H. Navarro, R. Martínez-Cuenca, G. Saavedra, M. Martínez-Corral y B. Javidi, "3D integral imaging display by smart pseudoscopic-to-orthoscopic conversion (SPOC)", *Opt. Express*, vol. 18, n.º 25, págs. 25 573-25 583, 2010 (vid. págs. 75, 76, 80).
- [76]H. Deng, Q. Wang y D. Li, "Method of generating orthoscopic elemental image array from sparse camera array", *Chinese Opt. Lett.*, vol. 10, n.º 6, págs. 1-3, 2012 (vid. pág. 75).
- [77]A. Dorado, G. Saavedra, J. Sola-pikabea y M. Martínez-corral, "Integral Imaging Monitors with an Enlarged Viewing Angle", *J. Inf. Commun. Conver. Eng.*, vol. 13, n.º 2, págs. 132-138, 2015 (vid. págs. 89, 90).
- [78]A. Dorado, M. Martínez-Corral, G. Saavedra y S. Hong, "Computation and Display of 3D Movie from a Single Integral Photography", *J. Disp. Technol*, vol. 12, n.º 7, págs. 343-348, 2016 (vid. págs. 89, 99, 102).
- [79]B. Lee, S. Jung y J.-H. Park, "Viewing-angle-enhanced integral imaging by lens switching", *Opt. Lett.*, vol. 27, n.º 10, pág. 818, 2002 (vid. pág. 89).
- [80]S. Jung, J.-H. Park, H. Choi y B. Lee, "Viewing-angle-enhanced integral three-dimensional imaging along all directions without mechanical movement", *Opt. Express*, vol. 11, n.º 12, pág. 1346, 2003 (vid. pág. 89).
- [81]J.-S. Jang y B. Javidi, "Improvement of viewing angle in integral imaging by use of moving lenslet arrays with low fill factor", *Appl. Opt.*, vol. 42, n.º 11, pág. 1996, 2003 (vid. pág. 89).

- [82]R. Martínez-Cuenca, H. Navarro, G. Saavedra, B. Javidi y M. Martínez-Corral, "Enhanced viewing-angle integral imaging by multiple-axis telecentric relay system.", *Opt. Express*, vol. 15, n.º 24, págs. 16 255-16 260, 2007 (vid. pág. 89).
- [83]M. Miura, J. Arai, M. Okui y F. Okano, "Method of enlarging horizontal viewing zone in integral imaging", *Three-Dimensional Imaging, Visualization, and Display*, vol. 8043, págs. 1-8, 2011 (vid. págs. 89, 94).
- [84]M. Miura, J. Arai, T. Mishina, M. Okui y F. Okano, "Integral imaging system with enlarged horizontal viewing angle", *Proc. SPIE*, vol. 8384, 838400-838400-9, 2012 (vid. págs. 89, 90, 92, 93).
- [85]H. Deng, Q.-H. Wang, L. Li y D.-H. Li, "An integral-imaging three-dimensional display with wide viewing angle", *Journal of the Society for Information Display*, vol. 19, n.º 10, págs. 679-684, 2011 (vid. pág. 89).
- [86]C. W. Chen, M. Cho, Y. P. Huang y B. Javidi, "Improved viewing zones for projection type integral imaging 3D display using adaptive liquid crystal prism array", *J. Disp. Technol*, vol. 10, n.º 3, págs. 198-203, 2014 (vid. pág. 89).
- [87]A. Dorado, S. Hong, G. Saavedra y col., "Toward 3D integral-imaging broadcast with increased viewing angle and parallax", *Optics and Lasers in Engineering*, vol. 107, págs. 83-90, 2018 (vid. pág. 109).
- [88]G. Park, J. Hong, Y. Kim y B. Lee, "Enhancement of viewing angle and viewing distance in integral imaging by head tracking", en *Advances in Imaging*, vol. 1, 2009, págs. 1-3 (vid. pág. 110).
- [89]G. Park, J.-H. Jung, K. Hong y col., "Multi-viewer tracking integral imaging system and its viewing zone analysis.", *Opt. Express*, vol. 17, n.º 20, págs. 17 895-17 908, 2009 (vid. pág. 110).
- [90]S. Hong, D. Shin, J.-J. Lee y B.-G. Lee, "Viewing Angle-Improved 3D Integral Imaging Display with Eye Tracking Sensor", *J. Inf. Commun. Conver. Eng.*, vol. 12, n.º 4, págs. 208-214, 2014 (vid. pág. 110).
- [91]S. Xin, M. Martínez-Corral y B. Javidi, "Head Tracking Three-Dimensional Integral Imaging Conversion", *J. Disp. Technol*, vol. 12, n.º 6, págs. 542-548, 2016 (vid. pág. 110).
- [92]Z.-L. Xiong, Q.-H. Wang, S.-L. Li, H. Deng y C.-C. Ji, "Partially-overlapped viewing zone based integral imaging system with super wide viewing angle", *Opt. Express*, vol. 22, n.º 19, pág. 22 268, 2014 (vid. pág. 110).

- [93]*Kinect for Xbox one components and specifications, 2013.* (Vid. pág. 117).
- [94]*Smeenk R. Kinect v1 and Kinect v2 fields of view compared* (vid. pág. 117).
- [95]D. Pagliari y L. Pinto, “Calibration of Kinect for Xbox One and Comparison between the Two Generations of Microsoft Sensors”, *Sensors*, vol. 15, n.º 11, págs. 27 569-27 589, 2015 (vid. pág. 117).